2017

A structural biology approach to the problem of antibiotic resistance in bacteria

Jared Armand Delmar
Iowa State University

A structural biology approach to the problem of antibiotic resistance in bacteria

by

Jared Armand Delmar

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Physics

Program of Study Committee:
James W. Evans, Major Professor
James Cochran
Kai-Ming Ho
Vincenzo Venditti
Jigang Wang

Iowa State University

Ames, Iowa

2017

EPIGRAPH

"A man may do an immense deal of good, if he does not care who gets the credit for it."

-Father Strickland, 1863

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to offer special thanks to my labmates Abhi, Nitin, and Eric, for countless inspiring research talks. Angie, Jani, Sylvia, Feng, and Jack, who taught me everything about our lab. David Sr., David Jr., and Sekhar, for their invaluable opinions. Ed, for always encouraging me to think independently. My parents, for their unconditional support. My girlfriend Erika, who got my six. My family, friends, and colleagues, too important and too many to name. The physics department faculty and staff, especially Lori, Gloria, Deb, and Linda, for their unyielding kindness. My committee members, for their patience. Anyone who has ever made me a cup of coffee. Without all of you, not only would this work be impossible, but it would be irrelevant.

ABSTRACT

X-ray crystallography remains the most robust method to determine protein structure at the atomic level. We demonstrate how these structural studies can directly contribute to unsolved problems in biology, with a focus on the growing problem of antibiotic resistance in bacterial infections. Multi-drug efflux transporters are common and powerful resistance mechanisms that are capable of extruding a number of structurally unrelated antimicrobials, including antibiotics and toxic heavy metal ions, from the bacterial cell. We begin by presenting the crystal structures of the individual pump components of the *Escherichia coli* Cus system, a paradigm for efflux machinery, and speculate on how these pumps assemble to fight diverse antimicrobials. In *Mycobacterium tuberculosis*, the cell wall is critical to the virulence and antimicrobial resistance of these pathogens. Recent work shows that the MmpL transporter family contributes to cell wall biosynthesis by exporting fatty acids and lipidic elements of the cell wall. The expression of the *M. tuberculosis* MmpL proteins is controlled by a complex regulatory network, including the TetR family transcriptional regulators Rv3249c and Rv1816. We demonstrate how the structures of these two proteins enhance understanding of the MmpL family of proteins and to develop new antibacterial tools to fight tuberculosis. *Neisseria gonorrhoeae* is a Gram-negative human pathogen and the cause of the STD gonorrhea. In *N. gonorrhoeae,* the MtrCDE multidrug efflux system mediates resistance to diverse antibiotics, nonionic detergents, antibacterial peptides, bile salts, and steroidal hormones. We have developed several techniques to assemble the complete MtrCDE tripartite efflux complex, which we present here. These efforts have culminated in a low-resolution structure of the bipartite MtrCD complex.

Finally, we apply our crystallography techniques to the problem of chloroplast cell division. In plants and algae, chloroplast division proceeds by binary fission, involving the coordinated assembly of four rings, both inside and outside the cell. We have determined the first high-resolution crystal structure of the *Arabidopsis thaliana* cell division protein PARC6. In addition, we obtained the co-crystal structure of PARC6 and PDV1, another protein within this network, revealing the molecular details of the intermembrane space interaction during chloroplast cell division.

SUMMARY

Chapter I

X-ray crystallography remains the most robust method to determine protein structure at the atomic level. However, the bottlenecks of protein expression and purification often discourage further study. Based on our experiences in expressing and purifying antimicrobial efflux proteins, we explain how a pure and homogenous protein sample can be successfully crystallized by the vapor diffusion method. Further, we demonstrate how these structural studies can directly contribute to unsolved problems in biology; specifically, the growing problem of antibiotic resistance in bacterial infections, a leading cause of death worldwide.

Chapter II

A common and powerful resistance mechanism, bacterial multi-drug efflux transporters are capable of extruding and detoxifying a number of structurally unrelated antimicrobials, including antibiotics and toxic heavy metal ions, facilitating bacterial survival in noxious environments. Those transporters belonging to the resistance-nodulation-cell division (RND) superfamily typically assemble as tripartite efflux complexes, spanning the inner and outer membranes of the cell envelope. In *Escherichia coli*, the CusCFBA complex, which mediates resistance to copper(I) and silver(I) ions, is the only known RND transporter with a specificity for heavy metals. Here, we present the crystal structures of the individual pump components of the Cus system, a paradigm for efflux machinery, and speculate on how RND pumps assemble to fight diverse antimicrobials.

Chapter III

The *Mycobacterium tuberculosis* cell wall is critical to the virulence and antimicrobial resistance of these pathogens. Recent work shows that the MmpL transporter family, which belongs to the RND superfamily, contributes to cell wall biosynthesis by exporting fatty acids and lipidic elements of the cell wall. The expression of the *M. tuberculosis* MmpL proteins is controlled by a complex regulatory network, including the TetR family transcriptional regulators Rv3249c and Rv1816. We have determined the high-resolution crystal structures of these two regulators. Combined with functional studies, the structures of these regulators give us clues on how they function, which is essential to understanding the mechanisms of the MmpL family of proteins and to develop new antibacterial tools to fight tuberculosis.

Chapter IV

Neisseria gonorrhoeae is a Gram-negative human pathogen and the cause of the STD gonorrhea. Recently, treatment of gonorrhea has been severely compromised by the development of antibiotic-resistance strains. The best characterized RND multidrug efflux system in *N. gonorrhoeae* is the MtrCDE multidrug efflux system. This tripartite complex mediates the export of hydrophobic antimicrobial agents, such as antibiotics, nonionic detergents, certain antibacterial peptides, bile salts, and gonadal steroidal hormones. The availability of the three-dimensional structures of these efflux transporters and their accessory proteins should allow us to block their function. However, this approach necessitates the high-resolution structure of a completely assembled RND tripartite efflux

complex. Our efforts to elucidate the crystal structure of the MtrCDE pump has culminated in a low-resolution structure of the bipartite MtrCD complex.

Chapter V

Finally, we apply our crystallography techniques to the problem of chloroplast cell division. In plants and algae, as in their cyanobacterial ancestors, chloroplast division proceeds by binary fission. The fission process involves the coordinated assembly of four rings both inside and outside the cell at the midplastid: the inner PD ring, the inner Z-ring, the outer PD ring, and the outer DRP5B ring. Within this network, a pair of paralogous inner envelope membrane proteins, PARC6 and ARC6, and a pair of paralogous outer envelope membrane proteins, PDV1 and PDV2, are likely responsible for bridging both the inner and outer membranes to connect the Z-ring and DRP5B ring. We have determined the first high-resolution crystal structure of the PARC6 protein. In addition, we obtained the crystal structure of this PARC6 domain in complex with the intermembrane space domain of *Arabiposis thaliana* PDV1, revealing the molecular details of the intermembrane space interaction during chloroplast cell division.

CHAPTER I

CRYSTALLIZATION OF MEMBRANE PROTEINS BY VAPOR DIFFUSION

Delmar, JA, Bolla, JR, Su, CC, and Yu, EW. (2015) *Methods Enyzmol.* 557: 363-92

Introduction

As of this writing, nearly 100,000 protein structures have been deposited to the PDB. Approximately 90% were solved by X-ray crystallography—by far the most successful technique used to study protein structure. However, of more than 30,000 unique X-ray structures, only approximately 1% represent membrane proteins (http://blanco.biomol.uci.edu/mpstruc/). In contrast to their share of the known structures, membrane proteins are estimated to comprise approximately 30% of all genes.[1,2] They play the critical roles of gatekeepers, mediating messages and materials moving into and out of the cell. For example, G-protein coupled receptors, which are responsible for vision, taste, smell and immune response in eukaryotes;[3,4] multidrug efflux transporters, which confer antibiotic resistance to pathogenic bacteria;[5] voltage-gated ion channels, such as those responsible for neuron firing and muscle contraction;[6,7] and porins, through which bacteria, mitochondria, and chloroplasts exchange material, such as nutrients, with the extracellular environment.[8,9] These proteins comprise at least 60% of all drug targets.[10,11]

Unfortunately, understanding the structures and action mechanisms of important integral membrane proteins is often hampered by the difficulties associated with expression, purification, and crystallization. There are at least two non-trivial tasks in the way of straightforward application of X-ray crystallography to membrane proteins: i) obtaining a sufficient quantity of purified protein and ii) producing high quality crystals.[12] The latter will

be the focus of this chapter. However, as any practicing membrane protein crystallographer can attest to, obtaining diffraction-quality crystals can be a formidable task, which requires a pure and homogeneous protein sample. In this respect, sample preparation and crystallization are inseparable. Thus, a significant portion of this review will be devoted to the topics of membrane protein expression and purification. In addition, extensive examples and detailed discussions, based on our experiences in expressing, purifying, and crystallizing antimicrobial efflux proteins, will be provided. Hopefully, by introducing the available methodologies and protocols, a rational approach can be made to the often anecdotal process of membrane protein crystallization.

Membrane Protein Expression

With few exceptions, membrane proteins are not abundant in their native environment. For example, *Escherichia coli* cells were found to contain only eight transcripts per cell corresponding to the inner membrane copper transporter CusA, even after induction.[13] Those few exceptions include the rhodopsins, porins, ATPases, photosynthetic reaction centers, and other light-harvesting complexes.[14] Coincidentally, the first membrane protein structure determined by X-ray crystallography was the photosynthetic reaction center of *Rhodopseudomonas viridis*.[15] It wasn't until 1998, more than a decade later, that a membrane protein structure was obtained from a non-native, *i.e.* recombinant, source.[16,17] To date, recombinant overexpression is the most common approach to obtain proteins of interest. Often, this step is the first bottleneck that must be overcome in order to produce the milligram quantities of pure protein necessary for a single crystallization experiment.[12,14]

Host

Thus far, *E. coli* remains the favorite host for heterologous protein expression. It grows quickly, tolerates a high cell density, and does so on a relatively lean diet.[18] The protein yield from *E. coli* cells is also particularly high, compared to other organisms. In one experiment, the expression of 20 membrane proteins in six host organisms (three prokaryotic, including *E. coli*, and three eukaryotic) was measured.[12] Overall, the highest yields were achieved with *E. coli* cells. Further, only two membrane proteins tested were not successfully overexpressed in *E. coli*.

Bacterial proteins do tend to express better in *E. coli*, in general, but it is not a cure-all for heterologous membrane protein expression. Especially for eukaryotic membrane proteins, other host organisms must be considered.[12,19,20] Often, eukaryotic proteins contain post-translational modifications, such as phosphorylation, glycosylation, methylation, or acetylation, which a non-native host may be incapable of reproducing.[21] *E. coli* and other bacterial hosts, in particular, tend to lack these post-translational capabilities.[18]

Translational machinery will also vary between the host and native organism. *E. coli* cells exhibit a clear codon bias, with the population of tRNA approximately proportional to the chromosomal frequency.[22] Occurrence of rare codons or many minor codons in the target protein sequence has a marked effect on expression levels. The reason for this is apparent when considering that an overexpressed membrane protein may represent a significant fraction of the total cellular protein. In some cases, even a single rare codon can be responsible for reduced expression levels.[22] Accordingly, optimal expression of a protein might be achieved by optimizing the codons for the particular host. This can be a time consuming process. For example, the *Neisseria gonorrhoeae* gene *norM*, encoding the

multidrug efflux pump NorM, contains nine rare codons for *E. coli*. To overexpress the *N. gonorrhoeae* NorM membrane protein in *E. coli* TOP10 cells, our lab had to spend a few months fixing each of these codons.[23] Even the *E. coli* gene *acrD*, which produces the aminoglycoside efflux pump AcrD, has 15 rare codons. Thus, it is incompatible with *E. coli* host cells. We had to correct each of these 15 codons in order to overexpress this membrane protein in *E. coli* BL21(DE3) cells. Fortunately, several companies, such as Integrated DNA Technologies (IDT), offer protein codon optimization tools to correct for the codon bias encountered in many host cells.

Upon translation, membrane proteins must then be chaperoned from the cytoplasm and properly inserted into the corresponding membrane. This is of particular importance to outer membrane proteins of Gram-negative bacteria. A proper signaling peptide is essential to guide the expressed protein across the inner membrane and anchor to the outer membrane. Differences in the lipid bilayer structure and composition, as well as differences in insertion machinery, of the native organism compared to the host determine whether this process results in an active membrane protein or an intractable aggregate.[24,25]

Finally, properly overexpressed, folded, and active membrane proteins tend to be toxic to their host.[12] The reason for this is obvious in the case of membrane channels or transporters, which directly control the permeability and delicate intracellular chemistry of their host, and can have direct roles in metabolism. In an experiment conducted with *Rhodopseudomonas blastica*, the overexpression of porins resulted directly in lysis of the host cell.[26,27] Thus, the amount of inducer, such as isopropyl β-D-1-thiogalactopyranoside or arabinose, as well as the expression time and temperature, are often found to be critical variables for optimal protein expression.

Various strains of *E. coli* have been commercially developed to address the challenges encountered in membrane protein overexpression. Among those used in our lab, the BL21(DE3) strain (Novagen) is deficient in OmpT and lon proteases, which allows for increased protein stability;[18] mutations in the *lac* promoter render C41(DE3) and C43(DE3) strains far less sensitive to the toxicity of membrane protein overexpression;[28,29] and the BL21-CodonPlus strain (Stratagen) contains additional copies of rare *E. coli* tRNAs.

In addition to using these specialized strains of *E. coli*, our lab also generates Δ*acrB* knockout strains by removing the *acrB* gene, which encodes the multidrug efflux pump AcrB. Deletions of abundant native membrane proteins have been shown to facilitate the purification process and produce a more homogenous protein.[27] In many cases, without this specific deletion, our attempts at crystallizing membrane proteins would result in crystals of AcrB, instead.

Vector

In general, it is not easy to predict whether an expression vector is suitable for the protein of interest. Choosing a vector that is compatible with both the membrane protein target and host organism, and one that is tailored to the desired expression and purification systems, is typically a trial-and-error process. Luckily, many powerful expression vectors are commercially available. In particular, our lab likes the pET-type vectors (Novagen) for the expression of α-helical transmembrane proteins. For β-barrel transmembrane proteins, we have inordinate success with the pBAD-type vectors (Life Technologies).

If a particular protein cannot be expressed with any combination of vectors and host cells, one may want to think about expressing the protein fused with green fluorescent

protein (GFP) or maltose-binding protein (MBP). In many cases, the expression levels of these fusion proteins are much higher than the parental ones.[20] It has also been reported that protein expression can be improved simply by single-point mutagenesis. To this end, Molina et al. subjected eight bacterial and one human membrane proteins to cycles of random mutagenesis.[29] In five of these proteins, expression levels could be increased, as much as 40-fold, by one round of mutagenesis corresponding to approximately 6 mutations per 1000 residues. Further, the structure of the first β_1-adrenergic receptor was solved with one mutation, cysteine-to-lysine, which reportedly improved expression.[30]

Membrane Protein Purification

Successfully overexpressed membrane proteins must be extracted and purified from the host's lipid membrane before crystallization, with the intention of providing as similar an environment as possible to the native bilayer (Fig. 1). At the same time, we wish to leave out all the properties that make the bilayer unsuitable for crystallization, such as heterogeneity, polarity, charge, etc.[24] Ideally, these should be replaced by the most important indicators of crystallizability: monodispersity, and stability.[31] This is a difficult task, indeed. As stated by Bowie, membrane proteins are "delicate weaklings, unable to withstand the rigors of life outside the safety of the bilayer."[32] While detergent-based solubilization is the most popular and effective technique to replace the bilayer, most membrane proteins are only marginally stable even in the best available detergent.[33]

The reason for their weakness is twofold: i) detergent, often smaller than the lipid environment it is replacing, may not completely cover the largely hydrophobic surface of the membrane protein and ii) membrane proteins, of diverse function, tend to exhibit much

conformational flexibility, with some conformations more stable than others. Without the ordering imposed by the bilayer, they are free to adopt those conformations that may favor aggregation.[33]

Detergent

Beginning with the first membrane protein crystal structure in 1980, detergents have been rationally designed especially for the purpose of facilitating protein purification and crystallization. The number of these detergents is now estimated in the 100s.[17] Nonetheless, only a small subset have been successful. The search for a suitable detergent benefits greatly from internet-based databases, which compile the crystallization conditions of all membrane proteins to date. For example, "Membrane Proteins of Known Structure" (http://*blanco.biomol.uci.edu/mpstruc/*) is particularly helpful.

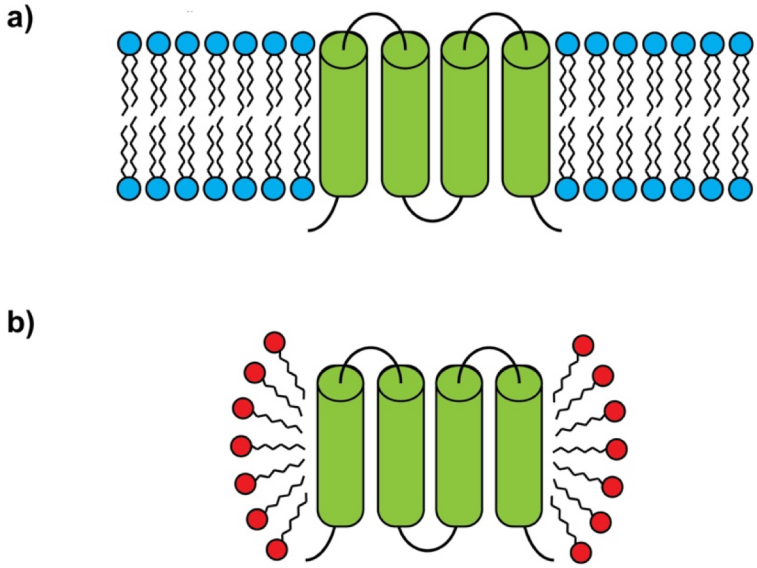

Figure 1. *Cartoon representation cross-section of a membrane protein solubilized in (a) lipid bilayer of the host cell and (b) detergent introduced during the purification process. Upon removal of the protein from the native lipid (blue) environment, the protein-detergent complex consists of a uniform disordered ring of detergent monomers (red) making hydrophobic contacts with the surface of the protein (green).*

At least structurally, all detergents are similar—a hydrophilic headgroup joined to a hydrophobic tail. This amphipathic character allows the detergent monomers to interact with both the hydrophobic surfaces of membrane proteins and solvent, as well as each other.[34] All

detergents used for membrane protein crystallization are alkyl-chain, between 7 and 14 carbons in length, with varying headgroups.[17] A detergent falls into one of three categories based on the properties of the headgroup: ionic, nonionic, or zwitterionic (Fig. 2).

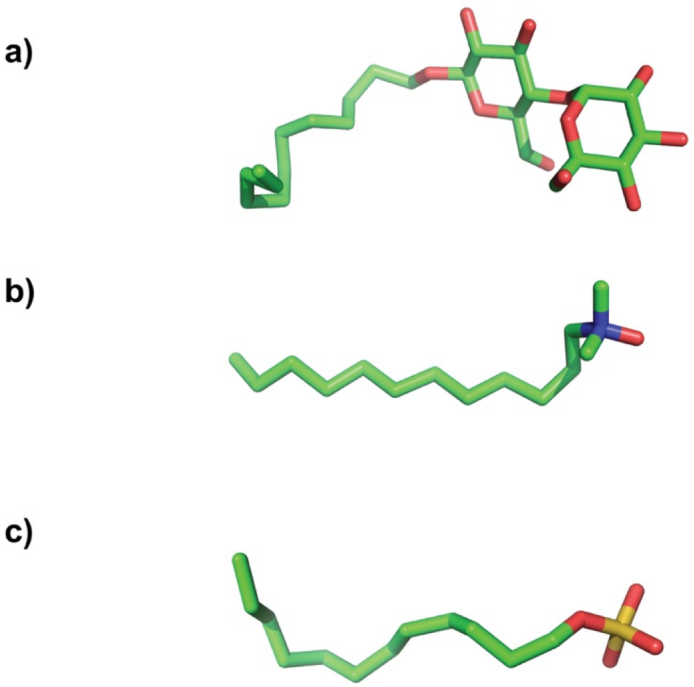

Figure 2. *Representative structures of commonly used detergents of three classes: (a) n-Dodecyl β-D-maltoside (DDM); non-ionic (b) Lauryldimethylamine-oxide (LDAO); zwitterionic. (c) Sodium dodecyl sulfate (SDS); anionic. Atoms are colored C, green; O, red; N, blue; and S, yellow.*

Importantly, the most effective families of mild detergents for purification and crystallization have non-ionic or, less commonly, zwitterionic headgroups. Besides solubilization, other detergent effects on membrane proteins can include inactivation, denaturation, or aggregation.[35] In these cases, harsh ionic detergents are often responsible. For example, the ionic detergent sodium dodecyl sulfate (SDS) is used in gel electrophoresis for exactly the purpose of protein denaturation. As of the publication of their paper in 2008, Newstead et al. reported that no outer membrane proteins to date had been crystallized with ionic detergent.[36]

To determine the effect of non-ionic and zwitterionic detergents on stability, Sonoda et al. used six common detergents to solubilize 24 prokaryotic and eukaryotic membrane

transporters.[37] The average stability measured was greatest for the two longest chain lengths of detergent, *n*-dodecyl-β-D-maltoside (DDM) and dodecyl nonaethylene glycol ether ($C_{12}E_9$), both non-ionic. In a similar study, Mancusso et al. examined 20 homologous lipid phosphatases.[31] All 20 purified membrane proteins tended to be unstable in DDM solution, and all precipitated when concentrated over 1 mg/mL. When DDM was exchanged for other detergents, both the stability and monodispersity of the sample varied. In one of the tested samples, LP-1, both these traits improved when solubilized with 6-cyclohexyl-1-hexyl-β-D-maltoside (CYMAL-6). The size-exclusion chromatography peaks were sharper and the protein could be concentrated to 4 mg/mL. Taken together, these results indicate the preference of membrane proteins for long-chain, non-ionic detergents. For instance, DDM, with its 12 alkyl-carbons, is well known to stabilize membrane proteins and is often used for both solubilization and crystallization.[38]

Protein-detergent complex

After purification in the best available detergent, the membrane protein carries between 40 and 200% of its own weight in detergent.[34,39] Save one or two molecules, all of these detergent coats are believed to be unstructured.[40] Their shapes are also known, based on the low-resolution neutron structures of two reaction centers, from *Rhodopseudomonas viridis* and *Rhodobacter sphaeroides*.[41,42] Remarkably, the shape of detergent surrounding each protein molecule is somewhat similar and is independent of the detergent molecules (*n*-dodecyl-*N*,*N*-dimethylamine-*N*-oxide (LDAO) and *n*-octyl-β-D-glucopyranoside (β-OG), respectively). In each case, the hydrophobic surface of the protein was covered in detergent molecules, wrapping the protein in a toroidal ring. Approximately 200 molecules cover 40%

of the surface area of each reaction center. The shear amount of detergent in each protein-detergent complex, the building blocks of the crystal lattice, demonstrates that the detergent is at least as important as the protein itself in crystallization.

Within the solvent surrounding the protein-detergent complex, detergent monomers exist which self-associate and bind to the membrane protein surface in a concentration-dependent manner, with a minimum concentration equal to the critical micelle concentration (CMC). For example, at 0.3% (w/v) β-OG, each monomer of human prostaglandin synthase was found to bind at least 40 molecules of detergent. At 0.7% (w/v) β-OG, the number of bound detergent molecules roughly doubled.[34] Crystallization is mostly done at detergent concentrations greater than the CMC, where equilibrium exists between detergent monomers, micelles, and detergent involved with the protein-detergent complex, with the concentration of monomers approximately equal to the CMC.[17] As a higher CMC detergent will have more monomers in solution, it is expected that those monomers will exchange more readily between detergent micelles and the protein molecule surface.[34]

Membrane Protein Crystallization via Vapor Diffusion

Many techniques exist to crystallize membrane proteins, including the recently developed lipidic cubic phase,[43,44] bicelle,[45] and vesicle fusion[46] methods. This chapter will focus on the vapor diffusion approach. The most common and successful method of membrane protein crystallization, to date, this approach has allowed researchers in our lab to determine the crystal structures of a number of membrane proteins. These include the inner membrane efflux pumps CusA,[47] AcrB,[48] and MtrD,[49] the outer membrane channels CusC,[50] MtrE,[51] and CmeC,[52] as well as the CusBA adaptor-transporter efflux complex.[53]

If not for the necessary presence of detergent, the crystallization of membrane proteins via vapor diffusion would be exactly similar to that of soluble proteins. While protein-protein interactions dominate the crystallization of soluble proteins, both detergent-detergent and protein-detergent interactions are additional considerations for the membrane protein crystallographer. Unlike crystallization of water-soluble proteins, choosing the right detergent has been the key success for our membrane protein crystallization efforts. In our opinion, this is the most important factor for obtaining high quality membrane protein crystals suitable for X-ray diffraction. Thus, the rule of thumb is that more efforts should be made to screen a variety of mild detergents than to optimize other parameters when crystallizing a new membrane protein.

In a simple crystallization experiment, a precipitant solution, containing at least buffer, precipitant, and salt components, is mixed with a protein solution, containing the purified protein-detergent complexes. This forces the protein solution into a state of supersaturation. To escape, the protein-detergent complex has two routes: i) aggregation and ii) crystallization. If the crystallization solution is not too harsh and the protein-detergent complex has enough time to make specific interactions with its neighbors, the latter results. In general, a slow introduction of the crystallization solution to the protein-detergent complex facilitates this, resulting in larger crystals with fewer imperfections.[54]

Compared to other techniques, the conditions for crystallizing membrane protein via vapor diffusion tend to be very mild, involving a mixture of relatively low concentrations of precipitant, salt, and buffer. In short, a small quantity of the purified protein solution is mixed, usually in equal ratio, with the crystallization solution contained in the well (precipitant solution). This drop, now half precipitant solution and half protein solution, is

either mounted in the well itself (sitting-drop) (Fig. 3a) or inverted above the well (hanging-drop) (Fig. 3b), and the chamber is made airtight. Vapor diffusion describes the evaporation of volatile species in the drop into the chamber, followed by diffusion across the chamber into the well. The end result of this process is chemical equilibrium between the drop and well. Because the well is much bigger, the equilibrium concentrations of drop components are approximated by the concentrations in the well.

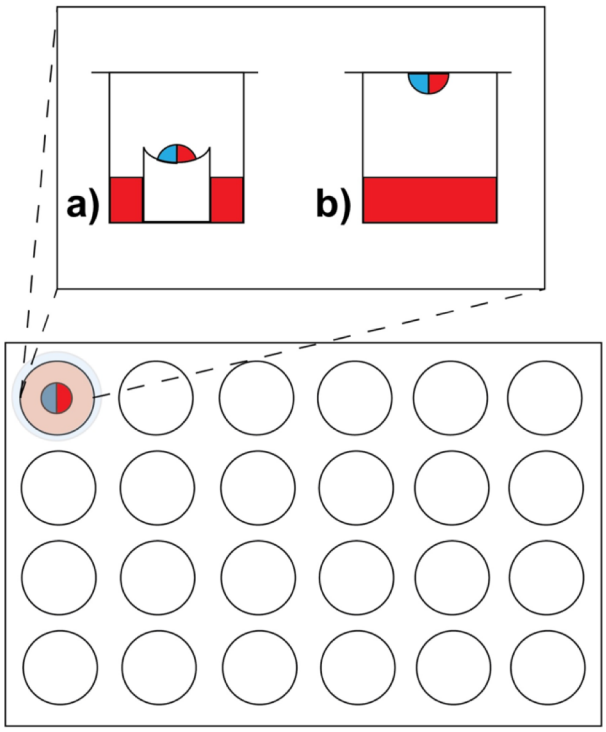

Figure 3. *A 24-well plate in which either (a) sitting-drop or (b) hanging-drop vapor diffusion is done. The well solution (colored red) is mixed in equal proportion with protein solution (colored blue). In the sitting-drop technique, the resulting mixture sits on a pedestal above the well solution and the top of the chamber is sealed with grease and glass. In the hanging-drop technique, the mixed drop is included on the glass slide itself, inverted and sealed with grease atop the chamber.*

For simplicity, it is generally assumed that only water is exchanged between the drop and well.[55] However, any volatile species can evaporate into the chamber and exchange between the drop and well. When it reaches equilibrium, or close to it, the hanging or sitting drop will manifest different types and degrees of phase behavior. In addition to the protein, the detergent and the

interactions between the two must be considered. The phase transition of detergent is relatively complex. It can crystallize, aggregate into micelles, or separate into detergent-rich, and detergent-poor phases.[25,56] Thus, searching for the right crystallization condition involves exploring point-by-point a complicated many-dimensional phase diagram for both protein and detergent.

The dimensions of this phase space include, at least, both protein and detergent concentrations, detergent type, precipitant concentration and type, buffer pH and type, salt concentration and type, and additive concentration and type. Further, the optimal crystallization conditions occupy only a very small volume of this phase space. Finally, the contribution of these factors to crystallization is non-linear. In general, the effect of one component on the solubility of the protein is coupled to the effect of another.[57] Even with the advent of robotic screening technology, capable of testing up to 40,000 conditions per day,[58] crystallization cannot be approached naively.

There are two general approaches to navigating the phase space of a protein: i) a 'grid search,' by systematic screening of what are believed to be the important variables,[59] or ii) using a 'coarse matrix screen,' heavily biased on known crystallization conditions of other proteins.[60] To date, many coarse matrix screens designed specifically for membrane proteins, including MemGold (Molecular Dimensions) and MembFac (Hampton Research), have been made commercially available. In many cases, these screening kits are instrumental in the search for initial crystallization conditions. The individual components of these conditions and their roles in the crystallization process will be discussed in the following sections, with an emphasis on the most common and successful ones.

Detergent

Some say that the biggest hurdle in membrane protein crystallization is not obtaining the crystals, but optimizing them for high-resolution X-ray diffraction.[37] In this respect, no other variable is more influential in the success of membrane protein crystallization than detergent.

Although no single detergent category can be generally applied to membrane protein crystallization, statistical evidence suggests that some are more useful than others. Almost all successfully used detergents are nonionic alkyl sugars with C7-C14 alkyl chains. By far, the most crystal structures have been solved using the alkylmaltoside family of detergents—nearly one half of all α-helical proteins.[61] This family includes the single most popular detergent, DDM, which contains 12 carbons in the main alkyl chain. By contrast, the alkylglucoside family of detergents are most commonly used to crystallize outer membrane proteins.[37] They include the second most popular detergent, β-OG, which is an 8 carbon alkyl-chained detergent.

Recently, the development of new detergents has emerged as a tool for handling difficult membrane proteins.[62] Non-conventional approaches, such as amphipols,[63] nanodiscs,[63] hemifluorinated surfactants,[63,64] cholate acid-based amphiphiles,[65] and the new family of maltose-neopentyl glycol amphiphiles[66] have been proven to aid in the solubilization of membrane proteins, while retaining their structure and function. This new generation of amphiphiles is particularly promising, and may eventually have a large role in the solubilization, stabilization, and crystallization of membrane proteins for high resolution structural analysis.

Precipitant

Based on their mechanism of action, precipitants can be divided into three categories: i) salts, ii) organic solvents, and iii) long-chain polymers. As will be discussed in the following section, salt acts as a precipitant by either dehydrating the protein molecules or altering the ionic strength.

Organic solvents, such as ethanol or isopropanol, generally act to reduce the dielectric constant of the crystallization solution. This makes electrostatic interactions weaker and reduces the solubility of ionic compounds slightly. Interestingly, one of the first membrane proteins to be crystallized, the hydrophobic seed protein crambin, was initially solubilized in ethanol and crystallized by slowly adding water.[34,67] At high concentrations, most of these compounds are denaturing. However, 2-Methyl-2,4-pentanediol (MPD) has been remarkably successful for outer membrane protein crystallization.[36]

The third and most successful category of precipitants comprises the long-chain polymers, including polyethylene glycols (PEGs). More than 80% of all membrane proteins have been successfully crystallized with various molecular weights of PEGs[36,61] and 100% of all membrane proteins in our lab. The mechanism of action of these polymers is attributed to volume exclusion effects.[57] Unlike other components of the crystallization solution, PEGs have no consistent conformation, resembling a flail. PEGs with molecular weights between 400 and 20000 Da have been effective for membrane protein crystallization, with the larger ones obviously being more forceful.

Salt

Salt ions, both cationic and anionic, can alter the solubility of proteins. This effect was first discovered by Franz Hofmeister.[68] The action mechanism is not entirely clear. However, it is thought to be due to specific interactions between ions, protein macromolecules, and water molecules in the solvent layer surrounding the protein. The effect of anions are usually ordered:

$$SO_4^{2-} > HPO_4^{2-} > CH_3CO_2^- > Cl^- > NO_3^- > Br^- > ClO_3^- > I^- > ClO_4^- > SCN^-$$

The order of cations is usually given as:

$$NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+}$$

Early members of the series tend to increase solvent surface tension and decrease the solubility of nonpolar molecules. They strengthen the hydrophobic interaction and this effect is often referred to "salting out". In this way, ammonium sulfate precipitation is commonly used in protein purification. Additionally, more than 10% of membrane protein crystals reported in the PDB are crystallized with the help of this salt.[61] By contrast, later ions in the series tend to increase the solubility of nonpolar molecules and decrease the order of water molecules. They weaken the hydrophobic interaction and this effect is called "salting in". Typically, ions that have a strong "salting in" effect are strong denaturants, because they interact much more strongly with the unfolded protein than with its native folded form.

Buffer

The actual net charge of the protein is determined not only by its composite amino acids, but by the pH of the solvent.[69] For each amino acid, its charge is determined by its pKa; at a pH below the pKa, an amino acid is positively charged and vice versa. Therefore,

altering the pH of the crystallization solution will alter the degree of electrostatic interaction, mainly, between the protein-detergent complexes.

At a pH equal to the pI, which is an average of the pKa of the surface residues of the protein, the net charge is neutralized and solubility drops dramatically. While this is usually too strong a condition for crystallization, it is common to find the optimal crystallization pH nearby.[69] The pH determines the charge of the zwitterionic detergents in a similar way.

Although no direct correlation appears to exist between predicted pI and the pH at which crystallization was reported, there is a good correlation between their difference, pH-pI, and pI. Specifically, acidic proteins tend to crystallize between 0 and 2.5 pH units above their pI, while basic proteins crystallize between 0.5 and 3 pH units below their pI.[70]

Case Studies

Crystallization of the E. coli CusA heavy-metal efflux pump

Initially, we cloned the *E. coli cusA* gene into pET15b to form the expression vector pET15bΩ*cusA*. This expression vector was then used to transform *E. coli* BL21(DE3) cells and overexpress the full-length CusA membrane protein, containing a 6xHis tag at the N-terminus. The expressed protein was purified using a Ni^{2+}-affinity column to >95% purity. The final purified protein was then dialyzed and concentrated to 20 mg/mL in buffer containing 20 mM Na-HEPES (pH 7.5) and 0.05% DDM for crystallization trials. The original protein crystals crystallized fairly quickly, and these crystals diffracted X-rays to 4 Å resolution. However, mass spectrometry identified these as crystals of the *E. coli* AcrB protein, which was probably co-purified with the CusA protein as a contaminant.

To avoid further contamination with the AcrB protein, we made an *E. coli* knockout strain BL21(DE3)*ΔacrB*, which harbors a deletion in the chromosomal *acrB* gene. The CusA transporter was then overexpressed using the transformed cells BL21(DE3)*ΔacrB*/pET15bΩ*cusA* and purified using a Ni^{2+}-affinity column. The purified protein was then dialyzed, concentrated to 20 mg/mL in buffer containing 20 mM Na-HEPES (pH 7.5) and 0.05% DDM, and subjected to crystallization trials using sitting-drop vapor diffusion. This time, mass spectrometry confirmed that the crystals were composed of CusA molecules. However, these crystals diffracted X-rays poorly, to ~9 Å resolution. After extensive screening, with as many salts, PEGs, pHs, temperatures, additives, and detergents available to us, we finally obtained plate-shaped crystals by adding 0.05% CYMAL-6 to the initial crystallization condition. Although these plate-shaped crystals were much better in quality, they diffracted X-rays anisotropically to a resolution of ~5 Å in one dimension and ~8 Å in the other.

As the addition of CYMAL-6 significantly improved the crystal quality, we decided to use CYMAL-6 to purify and crystallize the CusA protein. During this process, we also found that the additives Jeffamine M-600 (JM-600) and glycerol further improved their quality. Thus, the final crystallization condition was 2 μL protein solution (containing 20 mg/ml CusA protein in 20 mM Na-HEPES (pH 7.5) and 0.05% (w/v) CYMAL-6) mixed with 2 μL of reservoir solution containing 10% PEG 3350, 0.1 M Na-MES (pH 6.5), 0.4 M $(NH_4)_2SO_4$, 1% JM-600, and 10% glycerol. The resulting mixture was equilibrated against 500 μl of the reservoir solution. Typically, the dimensions of the crystals were 0.2 mm x 0.2 mm x 0.2 mm (Fig. 4a). Crystals of CusA were found to be trigonal with space group *R*32 ($a = b = 178.4$ Å, $c = 285.8$ Å) (Fig. 4b).[47] Based on the molecular weight of CusA (115.72

kDa), it was found that a single molecule occupied the asymmetric unit with a solvent content of 67.5%. The final crystal structure of the CusA transporter was determined to a resolution of 3.5 Å.

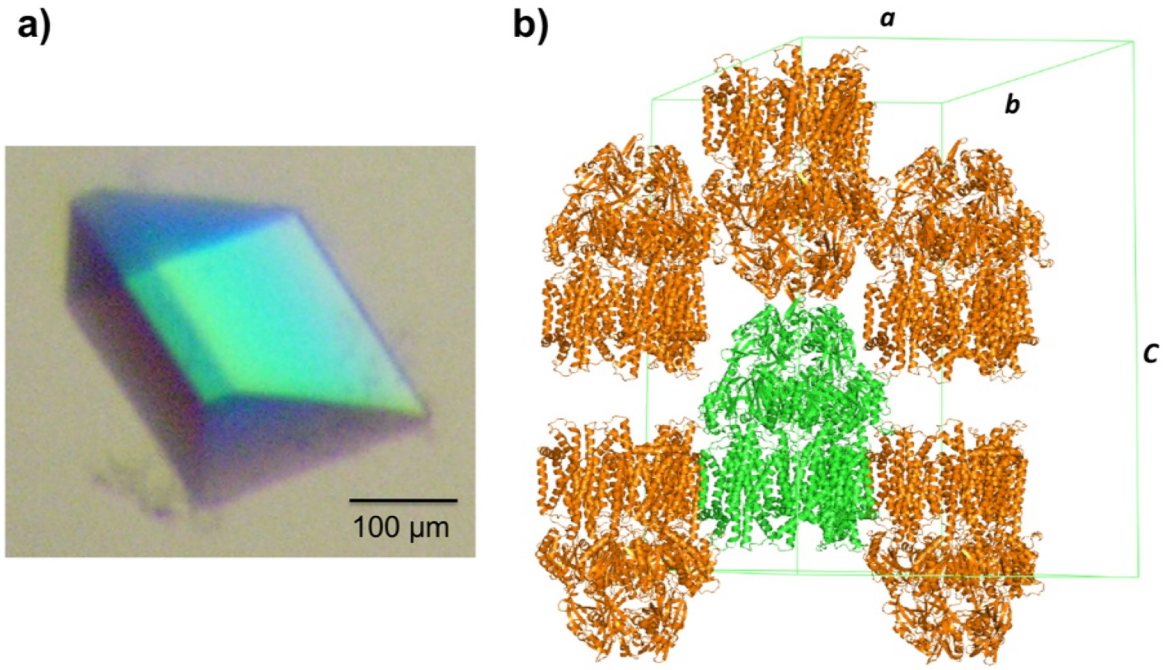

Figure 4. *Crystal structure of the E. coli inner membrane heavy metal efflux pump CusA. (a) A single crystal of CusA. (b) Packing diagram of the CusA crystal viewed orthogonal to the long axis of the unit cell.*

Crystallization of the E. coli CusBA adaptor-transporter complex

As of this article's publication, CusBA remains the only adaptor-transporter complex for which a high resolution X-ray structure is available. For purification of this important complex, the individual proteins of the CusA transporter and CusB[71] adaptor were individually overexpressed in *E. coli* cells BL21(DE3)*ΔacrB*/pET15bΩ*cusA* and

BL21(DE3)*ΔacrB*/pET15bΩ*cusB*, respectively. The CusA protein was solubilized with 2% CYMAL-6 and purified using Ni^{2+}-affinity chromatography, as mentioned previously. The CusB protein was also purified using a Ni^{2+}-affinity column, without the addition of any detergent. Co-crystals of the CusBA complex were obtained using sitting-drop vapor diffusion. After extensive optimization, the CusBA crystals were grown at room temperature in 24-well plates with a final crystallization condition of 2 μL protein solution (containing 0.1 mM CusA and 0.1 mM CusB in 20 mM Na-HEPES (pH 7.5) and 0.05% (w/v) CYMAL-6) mixed with 2 μL of reservoir solution containing 10% PEG 6000, 0.1 M Na-HEPES (pH 7.5), 0.1 M ammonium acetate, and 20% glycerol. The resulting mixture was equilibrated against 500 μL of the reservoir solution. Co-crystals of CusBA grew to a full size in the drops within two months. Typically, the dimensions of the crystals were 0.1 mm x 0.1 mm x 0.1 mm (Fig. 5a).

The co-crystals of the CusBA adaptor-transporter complex took a trigonal space group *R*32 with unit cell parameters $a = b = 160.2$ Å, $c = 682.7$ Å (Fig. 5b).[53] Based on the molecular weights of CusA (115.7 kDa) and CusB (42.3 kDa), the asymmetric unit contained one CusA and two CusB molecules with a solvent content of 70.8%. As it turns out, the presence of CusB drastically improved crystal quality in comparison with the crystals of CusA alone. The reason for this may be that CusB enlarges that hydrophilic surface of the membrane protein, thereby providing additional surface for crystal contacts. The final structure of the CusBA adaptor-transporter complex was determined to a resolution of 2.9 Å.

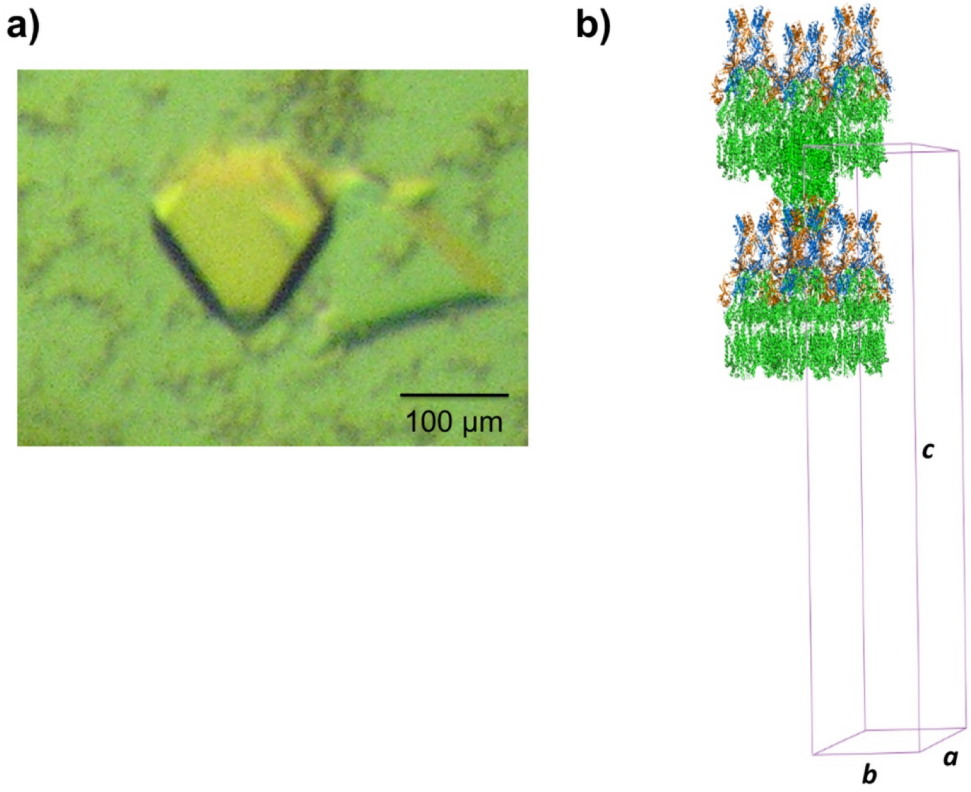

Figure 5. *Crystal structure of the E. coli inner membrane heavy metal efflux pump CusA in complex with the periplasmic membrane fusion protein CusB. (a) A single crystal of the CusBA complex. (b) Packing diagram of the CusBA cocrystal viewed orthogonal to the long axis of the unit cell.*

Crystallization of the N. gonorrhoeae MtrD multidrug efflux pump

It took us a few years of effort to crystallize the MtrD membrane protein. First, two different constructs, based on the protein sequence of MtrD in *N. gonorrhoeae* strain FA19, were used to express the MtrD protein, containing a 6xHis tag at the N- or C-terminus, respectively. The open reading frame (ORF) of *mtrD* from *N. gonorrhoeae* FA19 was cloned into pET15b to form each expression vectors. After screening of expression conditions, only the C-terminal 6xHis tagged MtrD protein could be expressed in *E. coli* C43(DE3)*ΔacrB*

cells, which harbor a deletion in the chromosomal *acrB* gene. Using these conditions, we then purified the MtrD protein using a Ni^{2+}-column. Optimization of the initial crystals with respect to primary and secondary detergent, salt, precipitant, pH and temperature, yielded only showers of microcrystals. Unfortunately, these microcrystals did not even diffract X-rays to low resolution.

Based on protein sequence alignments of different *N. gonorrhoeae* strains, we decided to focus on the homologous MtrD efflux pump from *N. gonorrhoeae* strain PID332. The alignment indicated that MtrD of strain PID332 is 11 amino acids shorter at the C-terminus in comparison with FA19 MtrD. We then cloned the *N. gonorrhoeae* PID332 *mtrD* gene into pET15b to generate the pET15bΩ*mtrD* expression vector. This recombinant plasmid encoding the *N. gonorrhoeae* PID332 MtrD protein, with a 6xHis at the C-terminus, was then transformed into C43(DE3)*ΔacrB* cells to overexpress this protein. Upon expression, MtrD was solubilized in 2% CYMAL-6 and subjected to purification using Ni^{2+}-affinity chromatography. Initial crystallization trials, using 0.05% CYMAL-6 as a primary detergent, did not yield any crystals. However, small cubic shaped crystals appeared in the drops after adding sucrose monododecanoate (SM) as a secondary detergent. Subsequent crystallization trials were carried out using both CYMAL-6 and SM detergents. After extensive optimization, the final MtrD crystals were grown at room temperature using sitting-drop vapor diffusion with the following procedures. A 2 μL protein solution containing 0.2 mM MtrD in buffer solution (20 mM Na-HEPES (pH 7.5), 0.05% (w/v) CYMAL-6, and 0.5% (w/v) SM) was mixed with 2 μL of reservoir solution, containing 30% PEG 400, 0.1 M Na-Bicene (pH 8.5), 0.1 M NH_4SO_4, 0.05 M $BaCl_2$, and 9% glycerol. The resulting mixture was equilibrated against 500 μL of the reservoir solution. Typically, the

dimensions of the crystals were 0.2 mm x 0.2 mm x 0.2 mm (Fig. 6a). Crystals of MtrD belonged to space group *R*32 ($a = b = 153.0$ Å, $c = 360.7$ Å) (Fig. 6b).[49] Analysis of the Matthews coefficient indicated the presence of one MtrD molecule (113.69 kDa) per asymmetric unit, with a solvent content of 66.7%. The final crystal structure of the MtrD transporter was determined to a resolution of 3.5 Å.

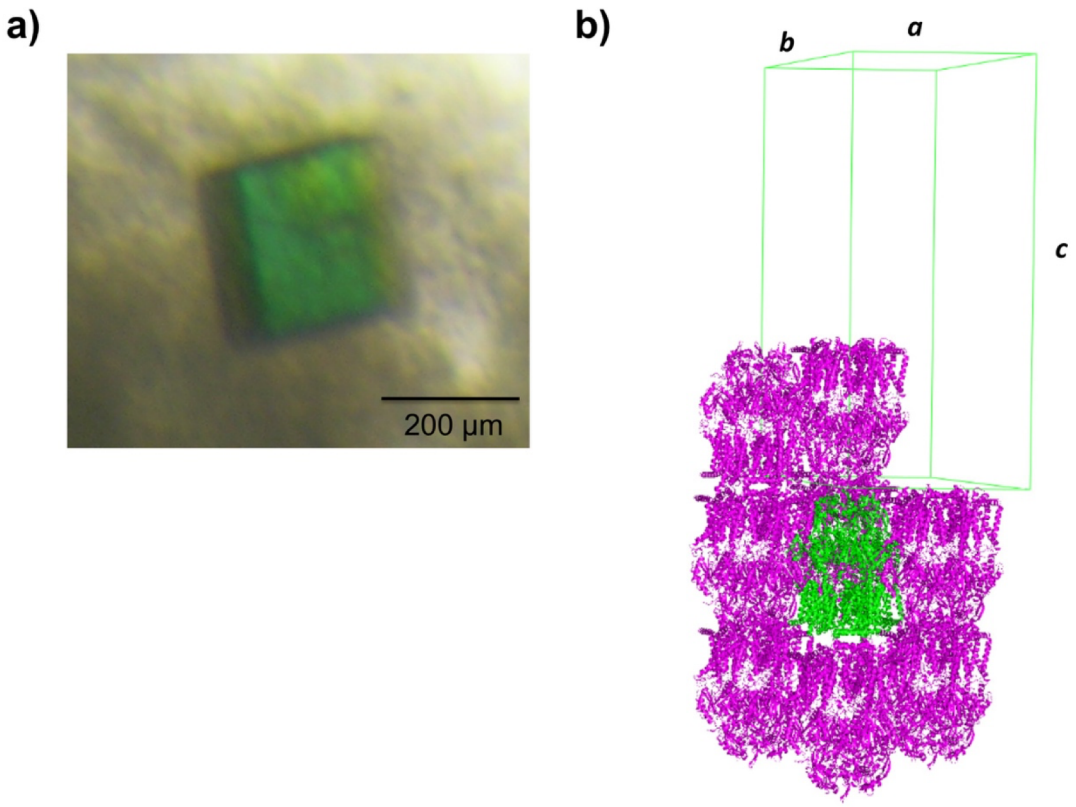

Figure 6. *Crystal structure of the N. gonorrhoeae inner membrane multidrug efflux pump MtrD. (a) A single crystal of MtrD. (b) Packing diagram of the MtrD crystal viewed orthogonal to the long axis of the unit cell.*

Crystallization of the E. coli CusC heavy-metal efflux channel

The CusC outer membrane channel is one of the most well-studied proteins in our laboratory. Its structures have allowed us to unmask the sequential transition of conformations leading to the folding and membrane insertion of this channel. To express CusC, the ORF of *cusC* from *E. coli* K12 genomic DNA was cloned into pBAD22 to produce the expression vector pBAD22Ω*cusC*. This expression system included a CusC signaling peptide at the N-terminus as well as a 6xHis tag at the C-terminal end. The N-terminal signaling peptide was needed to guide the channel protein to express at the outer membrane. It should be noted that this signaling peptide would be removed automatically when the CusC protein was matured in the cell. This recombinant plasmid was then used to transform the C43(DE3)*ΔacrB* cells to overexpress the channel protein. To ensure that the harvested CusC protein was attached to or anchored the *E. coli* outer membrane, and not the inner membrane, we performed a pre-extraction procedure using 0.5% sodium lauroyl sarcosinate to selectively dissolve and remove proteins of the inner membrane. The CusC outer membrane protein was then solubilized in 2 % (w/v) DDM. The extracted protein was purified with a Ni^{2+}-affinity column to >95% purity. The purified protein was then dialyzed and concentrated to 15 mg/mL in buffer containing 20 mM Na-HEPES (pH 7.5), 200 mM NaCl, and 0.05% (w/v) DDM for crystallization trials.

Sitting-drop vapor diffusion was employed for initial crystal screening. During this screening, we found that the full-length CusC channel could be crystallized only in the presence of 2% (w/v) β-OG, which served as a secondary detergent to DDM. It was also found that the addition of a small amount of JM 600 could improve the quality of the crystals. The best crystals of CusC were grown at room temperature with a reservoir solution

containing 8% PEG 3350, 0.05 M sodium acetate (pH 4.0), 0.2 M $(NH_4)_2SO_4$, 1% JM 600, and 2% β-OG. Typically, the dimensions of the crystals were 0.2 mm x 0.2 mm x 0.2 mm (Fig. 7a).

Crystals of the full-length CusC channel took the space group *R*32 with the unit-cell parameters $a = b = 88.5$ Å, $c = 474.42$ Å (Fig. 7b).[50] Based on the molecular weight of CusC (49.30 kDa), only one CusC molecule occupied the asymmetric unit with a solvent content of 67.8%. The final model was refined to a resolution of 2.1 Å.

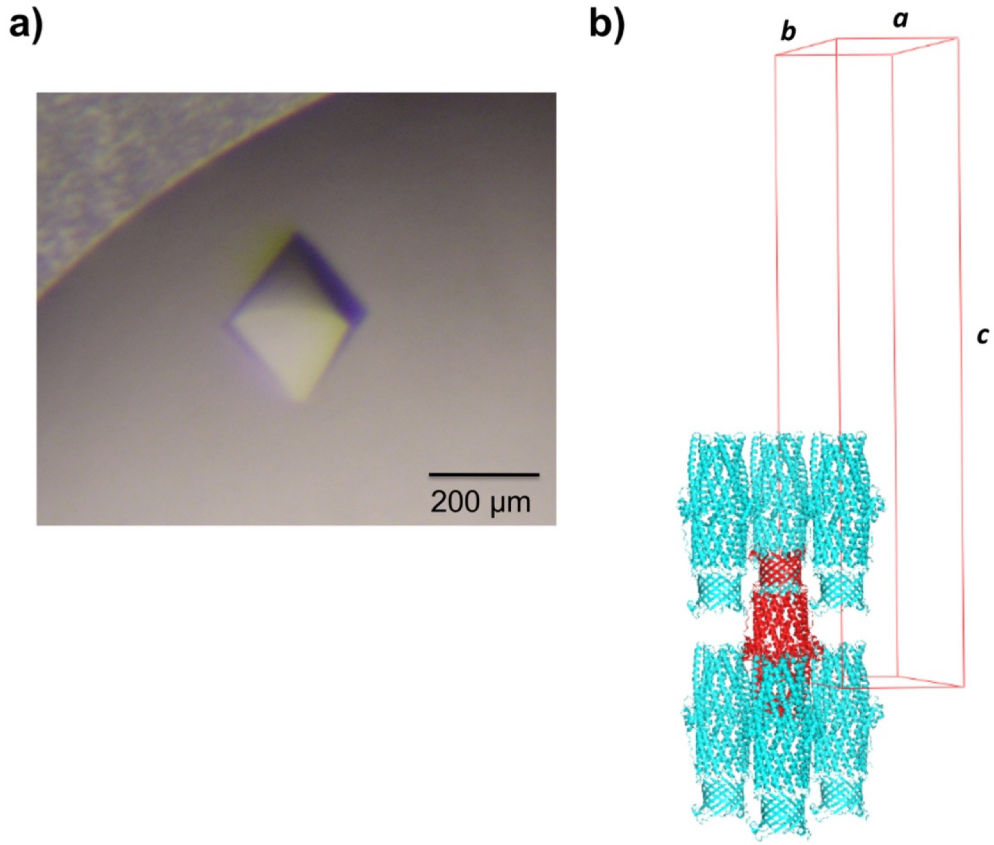

Figure 7. *Crystal structure of the E. coli outer membrane channel CusC. (a) A single crystal of CusC. (b) Packing diagram of the CusC crystal viewed orthogonal to the long axis of the unit cell.*

It is interesting to note that a single point mutation in the first cysteine residue of CusC resulted in a dramatically different conformation. Instead of the four-stranded beta-sheet present in the wild-type CusC monomer, the corresponding residues in this mutant form two independent random loops. By mutating or deleting this key residue, we obtained a different structure entirely, which corresponds to the state immediately before membrane insertion.[50] In this membrane protein, even the modification of one residue can prevent anchoring and insertion into the bilayer. Therefore, one should be cautious when crystallizing membrane proteins with any mutations or truncated sequences.

Crystallization of the N. gonorrhoeae MtrE multidrug efflux channel

The construct used to express the *N. gonorrhoeae* MtrE outer membrane channel was similar to that of *E. coli* CusC. However, after several trials, we found that the signaling peptide sequence of *Pseudomonas aeruginosa* OprM was the best peptide for expressing the *N. gonorrhoeae* MtrE heterogeneously in *E. coli*. In short, the ORF of *mtrE* from *N. gonorrhoeae* FA19 was cloned into pBAD22 to form the expression vector pBAD22Ω*mtrE*, which included the *P. aeruginosa* OprM signaling peptide at the N-terminus and a 6xHis tag at the C-terminus. The full-length MtrE protein was then expressed in *E. coli* C43(DE3) cells possessing this vector. The procedures for protein extraction and purification were the same as those for the CusC membrane protein. The purified protein was dialyzed and concentrated to 15 mg/mL in buffer containing 20 mM Na-HEPES (pH 7.5), 200 mM NaCl, and 0.05% (w/v) DDM for crystallization trials.

Crystals of MtrE were obtained using sitting-drop vapor diffusion. Based on crystal screening, the following procedure was adopted. A 2 μL protein solution containing 15

mg/ml MtrE protein in 20 mM Na-HEPES (pH 7.5), 200 mM NaCl, and 0.05% (w/v) DDM was mixed with a 2 μL of reservoir solution containing 20% PEG 400, 0.2 M sodium acetate (pH 4.6), 0.25 M $MgSO_4$, and 2% (w/v) β-OG. The resultant mixture was equilibrated against 500 μl of the reservoir solution at room temperature. Crystals of MtrE grew to a full size in the drops within two weeks. Typically, the dimensions of the crystals were 0.1 mm x 0.1 mm x 0.2 mm (Fig. 8a).

Crystals of the MtrE channel protein belonged to the space group $P6_322$ ($a = b = 93.9$ Å, $c = 391.5$ Å) (Fig. 8b).[51] Analysis of the Matthews coefficient indicated the presence of

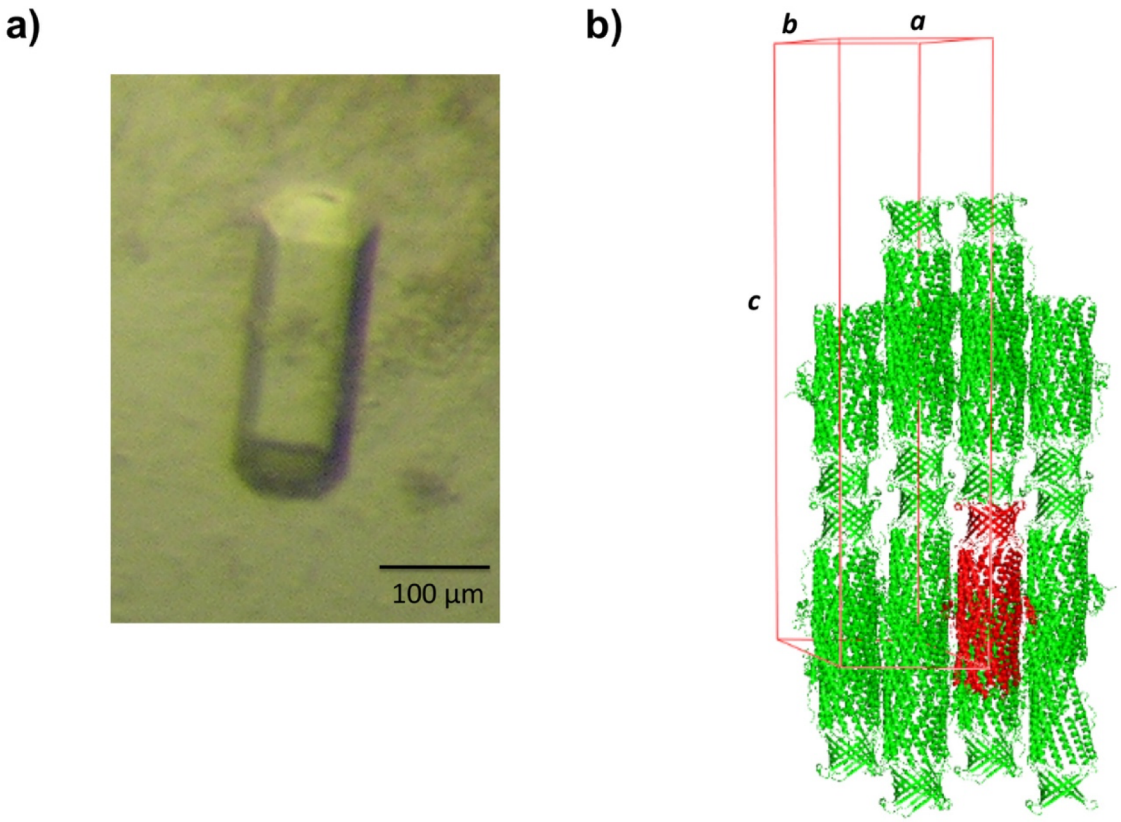

Figure 8. *Crystal structure of the N. gonorrhoeae outer membrane channel MtrE. (a) A single crystal of MtrE. (b) Packing diagram of the MtrE crystal viewed orthogonal to the long axis of the unit cell.*

one MtrE protomer (49.29 kDa) per asymmetric unit, with a solvent content of 75.8%. The final structural model was resolved to a resolution of 3.3 Å.

Crystallization of the C. jejuni CmeC multidrug efflux channel

The construct used to express the *C. jejuni* CmeC outer membrane channel was similar to that of *N. gonorrhoeae* MtrE. Briefly, the full-length CmeC membrane protein containing a 6xHis tag at the C-terminus was overproduced in *E. coli* C43(DE3)/pBAD22bΩ*cmeC* cells. This expression system included an OprM signaling peptide at the N-terminus and a 6xHis tag at the C-terminus. The purified protein was dialyzed and concentrated to 15 mg/mL in buffer containing 20 mM Na-HEPES (pH 7.5), 200 mM NaCl and 0.05% DDM.

Initial crystallization trials were not successful and did not yield even poor quality crystals. These trials were followed by extensive screening of secondary detergents. Fortunately, hexagonal shaped CmeC crystals were obtained, but only in drops containing the detergent C_8E_4. Further optimization of the crystallization conditions eventually produced well diffracting crystals.

Crystals of CmeC were obtained using sitting-drop vapor diffusion at room temperature in 24-well plates with the following procedures. A 2 μL protein solution containing 15 mg/mL CmeC protein in 20 mM Na-HEPES (pH 7.5), 200 mM NaCl, and 0.05% (w/v) DDM was mixed with 2 μl of reservoir solution containing 18% PEG 400, 0.1 M sodium acetate (pH 4.0), 0.3 M $(NH_4)_2SO_4$, and 2% C_8E_4. The resulting mixture was equilibrated against 500 μl of the reservoir solution. Crystals of CmeC grew to a full size in

the drops within two weeks. Typically, the dimensions of the crystals were 0.1 mm x 0.2 mm x 0.2 mm (Fig. 9a).

Crystals of the CmeC outer membrane channel belonged to the space group $C222_1$ with unit-cell parameters: a = 92.38 Å, b = 147.35 Å, c = 420.43 Å (Fig. 9b).[52] Based on the molecular weight of CmeC (54.26 kDa), three molecules were found in the asymmetric unit with a solvent content of 80.6%. These three molecules assembled to form a trimeric channel within the unit cell. The final crystal structure was determined to a resolution of 2.4 Å.

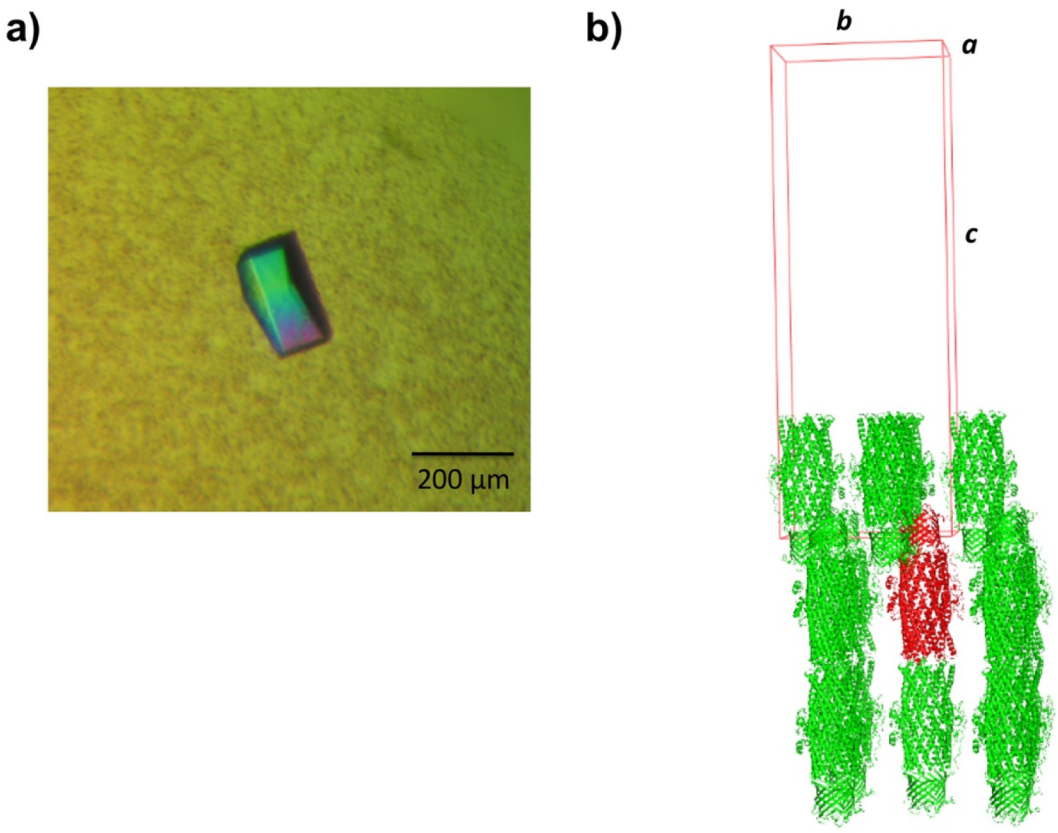

Figure 9. *Crystal structure of the C. jejuni outer membrane channel CmeC. (a) A single crystal of CmeC. (b) Packing diagram of the CmeC crystal viewed orthogonal to the long axis of the unit cell.*

Concluding Remarks

For the routine availability of three-dimensional membrane protein structures to become possible, there is still a mountain to climb. Like the state of soluble protein crystallization 20 years ago, initial data is still being acquired to facilitate the large-scale investigations demanded today. Only then can loftier goals of crystallization be achieved. For example, no atomic-resolution model of a drug efflux complex, spanning both the inner and outer bilayer, has been reported to date. In Gram-negative bacteria, such complexes are responsible for conferring resistance to commonly used antibiotics, a problem that is exacerbated by the increased use of these drugs. While bacterial infections remain a leading cause of death worldwide, invaluable structural information leading to antibiotic resistance mechanisms would provide a platform to produce new drugs and inhibitors that increase the efficacy of our weakened therapies.

References

1. Wallin, E., & von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*, *7*(4), 1029-1038.

2. Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. L. (2001). Predicting transmembrane topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology*, *305*, 567-580.

3. Lohse, M. J., Benovic, J. L., Codina, J., Caron, M. G., & Lefkowitz, R. J. (1990). Beta-Arrestin: a protein that regulates beta-adrenergic receptor function. *Science*, *248*(4962), 1547-1550.

4. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., & Miyano, M. (2000). Crystal structure of rhodopsin: a G protein-couple receptor. *Science*, *289*(5480), 739-745.

5. Saier, M. H., Jr., Tam, R., Reizer, A., & Reizer, J. (1994). Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport. *Molecular Microbiology*, *11*(5), 841-847.

6. Lee, S., Lee, A., Chen, J., & MacKinnon, R. (2005). Structure of the KvAP voltage-dependent K^+ channel and its dependence on the lipid membrane. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15441-15446.

7. Long, S. B., Tao, X., Campbell, E. B., & MacKinnon, R. (2007). Atomic structure of a voltage-dependent K^+ channel in a lipid membrane-like environment. *Nature*, *450*(7168), 376-382.

8. Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N., & Rosenbusch, J. P. (1992). Crystal structures explain functional properties of two E. coli porins. *Nature*, *358*(6389), 727-733.

9. Weiss, M. S., Abele, U., Weckesser, J., Welte, W., Schiltz, E., & Schulz, G. E. (1991). Molecular architecture and electrostatic properties of a bacterial porin. *Science*, *254*(5038), 1627-1630.

10. Drew, D., Fröderberg, L., Baars, L., & de Gier, J. L. (2003). Assembly and overexpression of membrane proteins in Escherichia coli. *Biochimica et Biophysica Acta*, *1610*(1), 3-10.

11. Overington, J. P., Al-Lazikani, B., & Hopkins, A. L. (2006). How many drug targets are there? *Nature Reviews Drug Discovery*, *5*(12), 993-996.

12. Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioiux, J., Salvi, D., Seigneurin-Berny, D., Richaud, P., Joyard, J., Pignol, D., Sabaty, M., Desnos, T., Pebay-Peyroula, E., Darrouzet, E., Vernet, T., & Rolland, N. (2011). Heterologous expression of membrane proteins: choosing the appropriate host. *PLoS ONE*, *6*, e29191.

13. Thieme, D., Neubauer, P., Nies, D. H., & Grass, G. (2008). Sandwich hybridization assay for sensitive detection of dynamic changes in mRNA transcript levels in crude Escherichia coli cell extracts in response to copper ions. *Applied and Environmental Microbiology*, *74*(24), 7463-7470.

14. Bill, R. M., Henderson, P. J. F., Iwata, S., Kunji, E. R. S., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C. G., & Vogel, H. (2011). Overcoming barriers to membrane protein structure determination. *Nature Biotechnology*, *29*(4), 335-340.

15. Deisenhofer, J., Epp, O., Miki, K., Huber, R., & Michel, H. (1985). Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3Å resolution. *Nature*, *318*(6047), 618-624.

16. Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., & MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K+ conductivity and selectivity. *Science*, *280*(5360), 69-77.

17. Wiener, M. C. (2004). A pedestrian guide to membrane protein crystallization. *Methods*, *34*(3), 364-372.

18. Sahdev, S., Khattar, S. K., & Saini, K. S. (2008). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry*, *307*(1), 249-264.

19. Schertler, G. F. X. (1992). Overproduction of membrane proteins. *Current Opinion in Structural Biology*, *2*(4), 534-544.

20. Pryor, K. D., & Leiting, B. (1997). High-level expression of soluble protein in Escherichia coli using a His6-tag and maltose-binding-protein double-affinity fusion system. *Protein Expression and Purification*, *10*(3), 309-319.

21. Gabrielsen, M., Gardiner, A. T., Fromme, P., & Cogdell, R. J. (2009). Membrane protein crystallization: approaching the problem and understanding the solutions. *Current Topics in Membranes*, *63*, 127-149.

22. Kane, J. F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Current Opinion in Biotechnology*, *6*(8), 494-500.

23. Long, F., Rouquette-Loughlin, C., Shafer, W. M., & Yu, E. W. (2008). Functional cloning and characterization of the multidrug efflux pumps NorM from Neisseria gonorrhoeae and YdhE from Escherichia coli. *Antimicrobial Agents and Chemotherapy*, *52*(9), 3052-3060.

24. Caffrey, M. (2003). Membrane protein crystallization. *Journal of Structural Biology*, *142*(1), 108-132.

25. Loll, P. J. (2003). Membrane protein structural biology: the high throughput challenge. *Journal of Structural Biology*, *142*(1), 144-153.

26. Schmid, B., Krömer, M., & Schulz, G. E. (1996). Expression of porin from Rhodopseudomonas blastica in Escherichia coli inclusion bodies and folding into exact native structure. *Federation of European Biochemical Societies Letters*, *381*(1), 111-114.

27. Bannwarth, M., & Schulz, G. E. (2003). The expression of outer membrane proteins for crystallization. *Biochimica et Biophysica Acta*, *1610*(1), 37-45.

28. Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Högbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O., & de Gier, J. (2008). *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14371-14376.

29. Molina, D. M., Cornvik, T., Eshaghi, S., Haeggström, J. Z., Nordlund, P., & Sabet, M. I. (2008). Engineering membrane protein overproduction in Escherichia coli. *Protein Science*, *17*, 673-680.

30. Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G. W., Tate, C. G., & Schertler, G. F. X. (2008). Structure of a β1-adrenergic G-protein-coupled receptor. *Nature*, *454*(7203), 486-491.

31. Mancusso, R., Karpowich, N., Czyzewski, B., & Wang, D. (2011). Simple screening method for improving membrane protein thermostability. *Methods*, *55*(4), 324-329.

32. Bowie, J. U. (2001). Stabilizing membrane proteins. *Current Opinion in Structural Biology*, *11*(4), 397-402.

33. Zhou, Y., & Bowie, J. U. (2000). Building a thermostable membrane protein. *Journal of Biological Chemistry*, *275*(10), 6975-6979.

34. Garavito, R. M., Picot, D., & Loll, P. J. (1996). Strategies for crystallizing membrane proteins. *Journal of Bioenergetics and Biomembranes*, *28*(1), 13-27.

35. De Grip, W. J. (1982). Thermal stability of rhodopsin and opsin in some novel detergents. *Methods in Enzymology*, *81*, 256-265.

36. Newstead, S., Hobbs, J., Jordan, D., Carpenter, E. P., & Iwata, S. (2008). Insights into outer membrane protein crystallisation. (2010). *Molecular Membrane Biology*, *25*(8), 631-638.

37. Sonoda, Y., Newstead, S., Hu, N., Alguel, Y., Niji, E., Beis, K., Yashiro, S., Lee, C., Leung, J., Cameron, A. D., Byrne, B., Iwata, S., & Drew, D. (2011). Benchmarking membrane protein detergent stability for improving throughput of high-resolution X-ray structures. *Structure*, *19*(1), 17-25.

38. Kang, H. J., Lee, C., & Drew, D. (2013). Breaking the barriers in membrane protein crystallography. *The International Journal of Biochemistry and Cell Biology*, *45*(3), 636-644.

39. Møller, J., & Le Maire, M. (1993). Detergent binding as a measure of hydrophobic surface area of integral membrane proteins. *Journal of Biological Chemistry*, *268*(25), 18659-18672.

40. Deisenhofer, J., & Michel, H. (1989). The photosynthetic reaction center from the purple bacterium Rhodopseudomonas viridis. *Science*, *245*(4925), 1463-1473.

41. Roth, M., Lewit-Bentley, A., Michel, H., Deisenhofer, J., Huber, R., & Oesterhelt, D. (1989). Detergent structure in crystals of a bacterial photosynethic reaction centre. *Nature*, *340*, 659-662.

42. Roth, M., Arnoux, B., Ducruix, A., & Reiss-Husson, F. (1991). Structure of the detergent phase and protein-detergent interactions in crystals of the wild-type (strain Y) Rhodobacter sphaeroides photochemical reaction center. *Biochemistry*, *30*(39), 9403-9413.

43. Landau, E. M., & Rosenbusch, J. P. (1996). Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(25), 14532-14535.

44. Nollert, P., Navarro, J., & Landau, E. M. (2002). Crystallization of membrane proteins in cubo. *Methods in Enzymology*, *343*, 183-199.

45. Faham, S., & Bowie, J. U. (2002). Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *Journal of Molecular Biology*. *316*(1), 1-6.

46. Takeda, K., Sato, H., Hino, T., Kono, M., Fukuda, K., Sakurai, I., Okada, T., & Kouyama, T. (1998). A novel three-dimensional crystal of bacteriorhodopsin obtained by successive fusion of the vesicular assemblies. *Journal of Molecular Biology*, *283*(2), 463-474.

47. Long, F., Su, C., Zimmermann, M. T., Boyken, S. E., Rajashankar, K. R., Jernigan, R. L., & Yu, E. W. (2010). Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport. *Nature*, *467*(7314), 484-488.

48. Su, C., Li, M., Gu, R., Takatsuka, Y., McDermott, G., Nikaido, H., & Yu, E. W. (2006). Conformation of the AcrB multidrug efflux pump in mutants of the putative proton relay pathway. *Journal of Bacteriology*, *188*(20), 7290-7296.

49. Bolla, J. R., Su, C., Do, S. V., Radhakrishnan, A., Kumar, N., Long, F., Chou, T., Delmar, J. A., Lei, H., Rajashankar, K. R., Shafer, W. M., & Yu, E. W. (2014). Crystal structure of the Neisseria gonorrhoeae MtrD inner membrane multidrug efflux pump. *PLoS ONE*, *9*(6), e97903.

50. Lei, H., Bolla, J. R., Bishop, N. R., Su, C., & Yu, E. W. (2014). Crystal structures of CusC review conformational changes accompanying folding and transmembrane channel formation. *Journal of Molecular Biology*, *426*(2), 403-411.

51. Lei, H., Chou, T., Su, C., Bolla, J. R., Kumar, N., Radhakrishnan, A., Long, F., Delmar, J. A., Do, S. V., Rajashankar, K. R., Shafer, W. M., & Yu, E. W. (2014). Crystal structure of the open state of the Neisseria gonorrhoeae MtrE outer membrane channel. *PLoS ONE*, *9*, e97475.

52. Su, C., Radhakrishnan, A., Kumar, N., Long, F., Bolla, J. R., Lei, H., Delmar, J. A., Do, S. V., Chou, T., Rajashankar, K. R., Zhang, Q., & Yu, E. W. (2014). Crystal structure of the Campylobacter jejuni CmeC outer membrane channel. *Protein Science*, *23*(7) 954-961.

53. Su, C., Long, F., Zimmermann, M. T., Rajashankar, K. R., Jernigan, R. L., & Yu, E. W. (2011). Crystal structure of the CusBA heavy-metal efflux complex of Escherichia coli. *Nature*, *470*(7335), 558-562.

54. Benvenuti, M., & Mangani, S. (2007). Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nature Protocols*, *2*(7), 1633-1651.

55. Fowlis, W. W., DeLucas, L. J., Twigg, P. J., Howard, S. B., Meehan, E. J., Jr., & Baird, J. K. (1988). Experimental and theoretical analysis of the rate of solvent equilibration in the hanging drop method of protein crystal growth. *Journal of Crystal Growth*, *90*, 117-129.

56. Newby, Z. E. R., O'Connell, J. D., III, Gruswitz, F., Hays, F. A., Harries, W. E. C., Harwood, I. M., Ho, J. D., Lee, J. K., Savage, D. F., Miercke, L. J. W., & Stroud, R. M. (2009). A general protocol for the crystallization of membrane proteins for X-ray structural determination. *Nature Protocols*, *4*(5), 619-637.

57. McPherson, A. (2004). Protein crystallization in the structural genomics era. *Journal of Structural and Functional Genomics*, *5*(1), 3-12.

58. Stevens, R. C. (2000). High-throughput protein crystallization. *Current Opinion in Structural Biology*, *10*, 558-563.

59. McPherson, A. (2009). Introduction to the crystallization of biological macromolecules. *Current Topics in Membranes*, *63*, 5-23

60. Jancarik, J., & Kim, S. H. (1991). Sparse matrix sampling: a screening method for the crystallization of proteins. *Journal of Applied Crystallography*, *24*, 409-411.

61. Newstead, S., Ferrandon, S., & Iwata, S. (2008). Rationalizing α-helical membrane protein crystallization. *Protein Science*, *17*(1), 466-472.

62. Bolla, J. R., Su, C., & Yu, E. W. (2012). Biomolecular membrane protein crystallization. *Philosophical Magazine (Abingdon)*, *92*(19), 2648-2661.

63. Popot, J. L. (2010). Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annual Review of Biochemistry*, *79*, 737-775.

64. Breyton, C., Gabel, F., Abla, M., Pierre, Y., Lebaupain, F., Durand, G., Popot, J. L., Ebel, C. & Pucci, B. (2009). Micellar and biochemical properties of (hemi)fluorinated surfactants are controlled by the size of the polar head. *Biophysical Journal*, *97*(4), 1077-1086.

65. Zhang, Q., Ma, X., Ward, A., Hong, W. X., Jaakola, V. P., Stevens, R. C., Finn, M. G., & Chang, G. (2007). Designing facial amphiphiles for the stabilization of integral membrane proteins. *Angewandte Chemie International Edition*, *46*(37), 7023-7025.

66. Chae, P. S., Rasmussen, S. G. F., Rana, R. R., Gotfryd, K., Chandra, R., Goren, M. A., Kruse, A. C., Nurva, S., Loland, S. J., Pierre, Y., Drew, D., Popot, J. L., Picot, D., Fox, B. G., Guan, L., Gether, U., Byrne, B., Kobilka, B., & Gellman, S. H. (2010). Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization, and crystallization of membrane proteins. *Nature Methods*, *7*(12), 1003-1008.

67. Teeter, M. (1984). Water structure of a hydrophobic protein at atomic resolution: pentagon rings of water molecules in crystals of crambin. *Proceedings of the National Academy of Sciences of the United States of America*, *81*(19), 6014-6018.

68. Hofmeister, F. (1888). Zur lehre von der wirkung der saize. *Archiv für Experimentelle Pathologie und Pharmakologie*, *24*(4), 247-260.

69. Rosenberger, F. (1996). Protein crystallization. *Journal of Crystal Growth*, *166*(1), 40-54.

70. Kantardjieff, K. A., & Rupp, B. (2004). Protein isoelectric point as a predictor for increased crystallization screening efficiency. *Bioinformatics*, 20(14), 2162-2168.

71. Su, C., Yang, F., Long, F., Reyon, D., Routh, M. D., Kuo, D. W., Mokhtari, A. K., Van Ornam, J. D., Rabe, K. L., Hoy, J. A., Lee, Y. J., Rajashankar, K. R., & Yu, E. W. (2009). Crystal structure of the membrane fusion protein CusB from Escherichia coli. *Journal of Molecular Biology*, 393(2), 342-355.

CHAPTER II

BACTERIAL MULTIDRUG EFFLUX TRANSPORTERS

Delmar, JA, Su, CC, and Yu, EW. (2014) *Annu. Rev. Biophys.* 43: 97-117

Multi-Drug Resistance in Bacteria

Antibiotics and other antimicrobial therapies have been used to treat infections for over a century. It is widely accepted that the increased use of these drugs has resulted in bacteria that are adapted and more resistant to the typical treatments. Until recently, the major means of antimicrobial resistance in bacteria was believed to be similar to the breakdown of penicillin by β-lactamase, i.e., modification of the offending compound by a drug-specific enzyme.[54,97] However, as antibiotic engineering advances, so too do the methods of resistance. Currently, there are three major mechanisms bacteria use to avoid the toxic effects of biocides.[80,97,102]

The first, enzymatic alteration, can be divided into two classes: 1) those enzymes that degrade the drug, like β-lactamase, and 2) those that chemically modify it. An example of the second class is the periplasmic copper oxidase CueO in *Escherichia coli*, which relieves the copper sensitivity of the cell by oxidizing toxic Cu(I) to less toxic Cu(II).[113,114] As metals cannot be chemically degraded, bacteria must employ alternative methods of detoxification in the case of copper or silver poisoning.[64,77] Accordingly, the second method of bacterial drug resistance involves enzymatic alteration of the target of the drug, rather than the drug itself. Fluoroquinolone resistance in bacteria is most commonly attributed to this mechanism; specifically, modification of the enzyme DNA gyrase.[41,97] While the first two methods are limited due to their specificity for a single drug or drug-binding site, the third mechanism is

more general. By expelling a toxin before it can act, or by preventing it from entering the cell altogether, a bacterium utilizes its most potent form of resistance.

The first line of cellular defense is either its surrounding peptidoglycan web (Gram-positive) or the outer lipid membrane (Gram-negative).[79] However, the peptidoglycan shell is too coarse to limit the diffusion of very small molecules. Thus, Gram-negative bacteria are inherently less susceptible to toxins.[81] *Pseudomonas aeruginosa* is one example of a pathogen with exceptionally high resistance to a wide range of antibiotics. It also possesses exceptionally narrow porins—outer-membrane proteins that allow passive diffusion of small molecules like sugars and amino acids. Thus, *P. aeruginosa* possesses an exceptionally-impermeable outer membrane.[4,80,81] Despite this fact, neither simple nor porin-facilitated diffusion across the cell membrane is enough to explain the intrinsic antimicrobial resistance of certain bacteria. In two related studies, the minimum inhibitory concentration (MIC) of antibiotics tetracycline and cephalothin in *E. coli* increased by only factors of 1.5 and 2, respectively, when porin expression was interrupted.[42,55,91] Furthermore, it is common for the most effective antibiotics to exhibit some hydrophobicity, which increases the permeability of the lipid bilayer to these drugs.[80]

Not diffusion, but active transport is now recognized as the major player in antimicrobial resistance.[37,97] To remove toxic molecules that enter via porins or porin-like transport systems, the cell utilizes a powerful pumping mechanism. Composed of one or more protein components, these multi-drug efflux pumps traverse the cell membrane, bind and actively pump out a broad range of noxious agents; in some cases, from the cytoplasm all the way to the outside of the cell. Metal-ion and antibiotic resistance systems of this type have been found encoded on plasmids of every eubacterial group tested, from *E. coli* to

Streptomyces sp.[109] For example, *E. coli*'s *pco* (for plasmid-borne copper resistance determinant) and *P. syringae*'s *cop* operon.[14] Many important resistance systems, including the subject of this review, are chromosomally represented, as well.

Heavy-Metal Resistance in Bacteria

In addition to preventing the action of drugs, bacteria utilize efflux pumps to regulate the delicate cellular levels of metal ions, which are among the most common enzymatic cofactors. Around 40% of all enzymes contain transition metals such as Mg, Zn, Fe, Mn, Ca, Co and Cu, and even Cd has been found in the carbonic anhydrase of a certain algae.[52,77,126] Cytochrome c oxidase, the last step in the mitochondrial and bacterial electron transport chain; prokaryotic superoxide dismutase, which detoxifies dangerous superoxide radicals; lysyl oxidase, a collagen-elastin cross-linker; and tyrosinase, a melanin producer, are all enzymes that require bound copper ion to function.[108] However, while transition metals such as copper are important and necessary components of the cellular environment, in excess they are extremely toxic. In recent experiments, *E. coli* cultures showed diminished to no growth at $CuCl_2$ concentrations as low as 1 mM.[8,76] and in another study growth defects were observed at concentrations exceeding 8 μM.[64] Non-essential metals can be toxic at even lower concentrations. Silver, for example, is such an efficient biocide that its effects can be observed on the nM scale.[98]

The primary cause of this toxicity is the high reactivity of transition metals; e.g., cations in particular have a high affinity for thiol and thioether residues, imidazoles, sulfides and nucleic acids.[24,77] Furthermore, most transition metals have similar binding affinities for these groups. Copper, however, is exceptional for its ability to push nearly any other metal

out of complex,[76] and a mismatch of enzyme and metal ion cofactor will typically result in inactivation of the enzyme. Cytoplasmic copper interferes with the formation of iron-sulfur clusters, necessary for the activity of, for example, the citric acid cycle enzyme aconitase or the heme group of cytochrome c.[76,77] Finally, due to their high redox potential, copper and iron are able to participate in the following Fenton-type reaction:

$$Cu^{+} + H_2O_2 \rightarrow Cu^{2+} + OH^{-} + OH$$

which generates dangerous hydroxyl radicals from hydrogen peroxide and Cu^{+} or Fe^{2+}.[34,64,77] These radicals are capable of reacting with basically anything and can cause damage to DNA, lipids and proteins.

While the mechanisms of heavy-metal toxicity have only been elucidated over the last century, the biocidal properties of copper and silver have been exploited for millennia.[18,29,111] The use of silver is increasingly common, and both copper and silver are popular and effective sterilizers, *viz.*, brass doorknobs, silverware, copper and silver water filters, silver-coated catheters, silver-containing bandages, mineral sanitizers in pools and silver nitrate and silver sulfadiazine, which are common antiseptics used to treat burns.[29,89,110,112] Many transition metal ions are essential cellular ingredients, however the efficacy of these metal-based antimicrobial treatments has forced bacteria to handle copper and silver with the same machinery as they do to antibiotics.

Classes of Multi-Drug Efflux Pumps

Based on sequence similarity, transport function, substrate specificity and energy coupling, ten families of efflux proteins have been classified as members of five superfamilies:[100,101] 1) the adenosine triphosphate (ATP)-binding cassette (ABC)

superfamily; 2) the major facilitator superfamily (MFS); 3) the small multi-drug resistance (SMR) superfamily; 4) the resistance-nodulation-cell division (RND) superfamily and 5) the multi-drug and toxic compound extrusion (MATE) superfamily (Fig. 1).[13,86,99,124] As an energy source, ABC transporters utilize ATP to drive toxins from the cell, while the other four superfamilies rely on an electrochemical gradient. Specifically, MFS, RND and SMR proteins employ the proton-motive force and the MATE superfamily is characterized by either Na^+- or H^+-substrate antiport.[102]

Gram-negative bacteria have been found to contain members of all five superfamilies,[82] which contribute to their intrinsic resistance to diverse antimicrobials. RND transporters are capable of forming powerful cooperative multi-protein structures that bridge

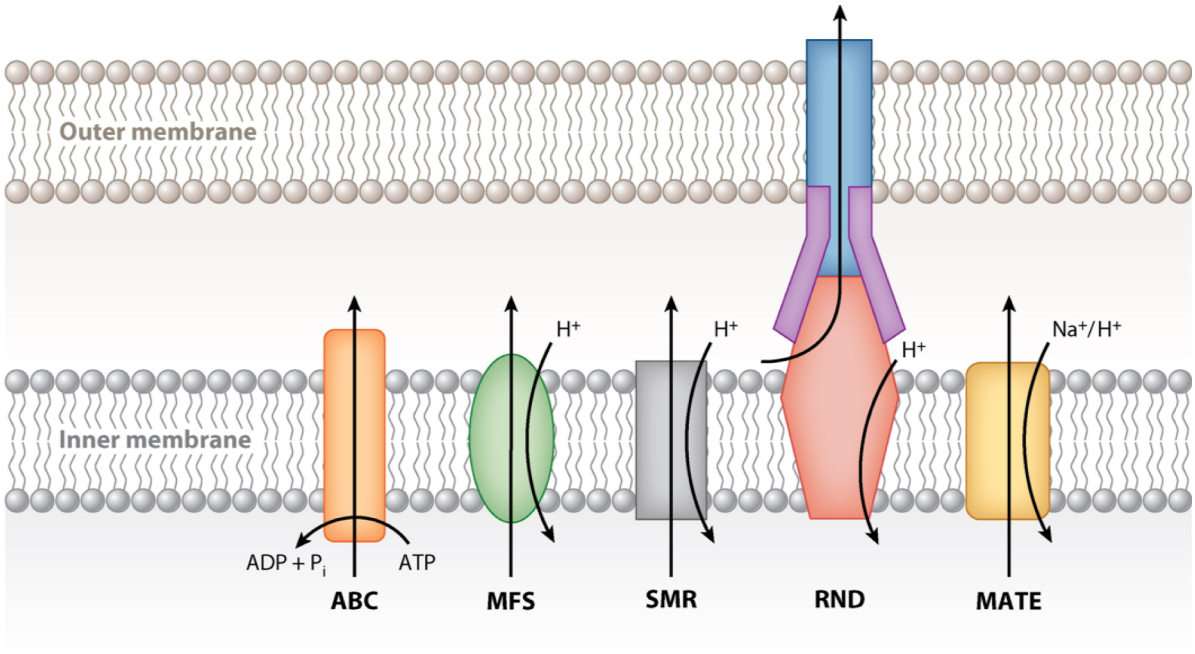

Figure 1. *Functional diversity among efflux proteins. Based on mode of transport, energy coupling mechanism, and phylogeny, these proteins are divided into five superfamilies. ABC superfamily proteins use ATP to transport diverse antimicrobials across the cellular inner membrane. MFS, SMR, and RND proteins function via an H^+-substrate antiport mechanism. MATE proteins use both H^+ and Na^+ as energy sources. RND transporters in particular can form multiprotein structures that bridge the inner and outer membranes.*

both the inner and outer membranes.[132] They are essential to the multi-drug resistance observed in many pathogens.[97]

Unfortunately, structural information is available for few members of the RND superfamily. Whereas only two have resolved crystal structures, including AcrB[71] and CusA,[62] there are a total of seven known RND proteins in *E. coli*. These seven transporters can be categorized into two distinct sub-families. AcrB,[21,70,71,83,84,90,103,105,116,123,130,131] AcrD,[1,84,96] AcrF,[53,63,84] MdtB,[10,47,72,84] MdtC[10,47,72,84] and YhiV[12,49,84] are multi-drug efflux pumps of the hydrophobic and amphiphilic efflux RND (HAE-RND) protein family[75,123] and CusA[62,84,117] belongs to the heavy-metal efflux RND (HME-RND) family.[75,123]

CusA, which specifically recognizes and confers resistance to Ag(I) and Cu(I) ions, is one of few known members of the HME-RND family, which includes CusA homolog SilA in *Salmonella typhimurium*[32] and CzcA of *Ralstonia sp.* CH34.[27,78] CusA is unique not only as the only known *E. coli* HME-RND transporter, but also as the only HME-RND protein, among all organisms, with available crystal structures. Among all RND efflux systems, CusCFBA is one of just three, including MexAB-OprM[2,3,38,106] and AcrAB-TolC,[50,67,104,123] with structures for each component. It is also the first efflux complex for which the adaptor and transporter were co-crystallized.[119]

Four proteins are found to operate in conjunction with CusA, making it only one of two known tetrapartite efflux systems. The fully-assembled complex consists of the inner-membrane, substrate-binding transporter, or pump, CusA; the periplasmic membrane fusion protein (MFP), or adaptor, CusB; the outer membrane factor (OMF), or channel, CusC; as well as the small, periplasmic, metal-binding protein, or chaperone, CusF, a component that has no analog in HAE-RND complexes (Fig. 2). While HAE-RND proteins are typically of

broad specificity, HME-RND proteins have been shown to be very specific, with the ability to distinguish between monovalent and divalent ions.[8,123]

Despite their differences, metal and antibiotic resistance systems are thought to have much in common.[9] In fact, the line that distinguishes one from another is sometimes blurred. In addition to copper and silver, the *cus* determinant has been implicated in *E. coli* resistance to drugs dinitrobenzene, dinitrophenol and fosmomycin.[19,84] Thus, both HAE- and HME-RND proteins are believed to be key components of the acquired and intrinsic resistance of Gram-negative pathogens to antimicrobials. The acquisition of even one RND system by a bacterium can increase its tolerance to a broad spectrum of drugs. For example, the Acr system in *E. coli* has been linked with resistance to tetracycline, chloramphenicol, fluoroquinolones, novobiocin, erythromycin, fusidic acid, rifampin, ethidium bromide, acriflavin, crystal violet, sodium dodecyl sulfate (SDS) and

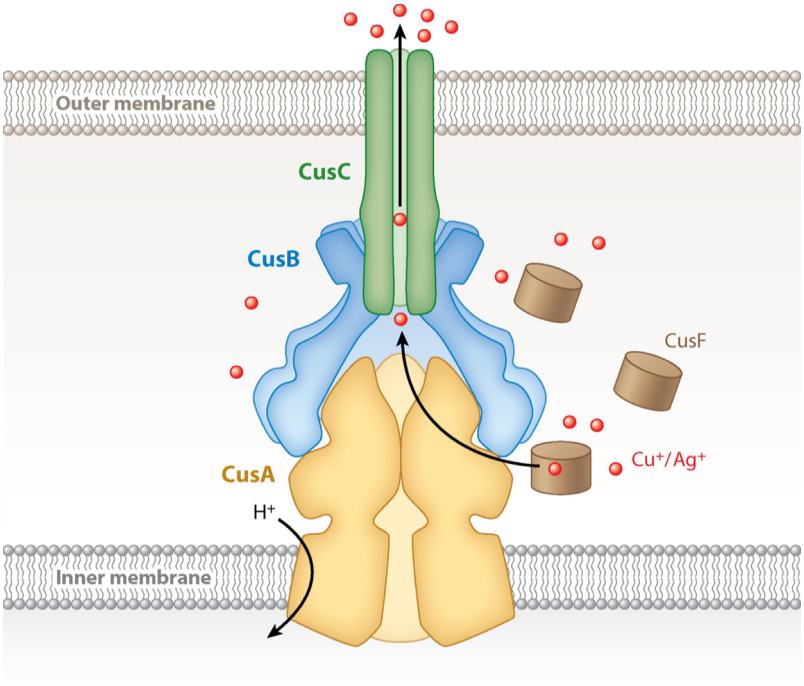

Figure 2. *Model of the fully assembled CusCFBA tetrapartite efflux system. In E. coli, this HME-RND complex is responsible for extruding Cu^+ and Ag^+ ions directly from the cell using the proton-motive force. The complex consists of the inner membrane transporter CusA (yellow), the membrane fusion adaptor CusB (blue), the outer membrane channel CusC (green), and the periplasmic metallochaperone CusF (purple). Cu^+ and Ag^+ ions have been found in complex with CusA, CusB, and CusF.*

deoxycholate—an unprecedented number of antibiotics, chemotherapeutics, detergents and dyes.[81] Perhaps as a direct consequence of the use of these compounds, including metal-based antimicrobial therapies, RND resistance systems are being discovered with increasing frequency.[48,93,110] Thus, these resistance systems pose a formidable clinical threat. As their prevalence in pathogens increases, so too does the need to understand them.

Although this review will focus on the details of the CusCFBA tetrapartite Cu(I)/Ag(I) transporter, many similarities are expected to be found between heavy-metal extrusion and antibiotic extrusion in general. Where appropriate throughout the article and as much as possible, parallels will be drawn between Cus and other known RND systems, with emphasis on the well-characterized MexAB-OprM and AcrAB-TolC efflux pumps, model systems for bacterial drug resistance.

Copper and Silver Resistance in E. coli

By the time the *E. coli* genome was completely sequenced,[11] already two HME-RND pumps and their corresponding complexes had been found in other organisms: CzcCBA in *Ralstonia* sp. CH34[27,73,78] and SilCBA in *S. typhimurium*.[32,110,112] The earliest example SilA, with nearly 87% sequence identity to CusA, was purified from a hospital strain of *S. typhimurium* at Massachusetts General in 1975. *S. typhimurium* expressing this plasmid-encoded transporter were found to be resistant to a “disturbing” number of antibiotics and metal ions, including the popular silver nitrate.[65] Ultimately, infection led to the deaths of three patients and the closing of the burn ward.[110] Based in part on their sequence similarity to the deadly *sil*, *czc* and other known antimicrobial efflux pumps, 37 open reading frames in

the *E. coli* genome were identified as putative drug transporters. In addition to CusA, this list includes six other RND proteins.[84,87]

The architecture of the *cus* locus consists of two back-to-back operons, *cusCFBA* and *cusRS* (formerly *ylcBCD-ybdE* and *ylcA-ybcZ,* respectively). The *cusA* gene is preceded by *cusB,* encoding the MFP, and *cusC,* encoding the OMF. Between *cusB* and *cusC* is the short ORF *cusF*, homologous to the corresponding gene in *sil,* for a small periplasmic chaperone protein.[25] Transcribed in the opposite direction are *cusR* and *cusS*, encoding a response regulator and histidine kinase, respectively. Upstream of *cusA* and *cusB* is a promoter region with high sequence similarity to those of *pcoAp*, *pcoEp*, *silCp*, *silEp*, *copAp* and *copHp*—all copper and silver resistance genes of Gram-negative bacteria.[25,32,78] Although the function of each *cus* gene was unknown at the time of discovery, all previously determined *cusA* homologs were found to be preceded by a *cusB* homolog and organized as an operon, including the *acr* and *mex* loci. In the closely-related heavy-metal resistance domains *czc* and *cnr* of *R. metallidurans* and the *sil* determinant of *Salmonella*, the OMF gene is also part of the operon. However, in the case of *acr* and *mex*, the outer membrane channel genes (*tolC* and *oprM*, respectively) were found elsewhere.

Several early experiments sought to characterize the new genes and their results led to the discovery that the *cus* locus encodes, primarily, a copper- and silver-transporting machine. This makes *cus* special in *E. coli* as one of only two chromosomally-encoded copper resistance systems and the only one implicated in silver resistance. The other copper locus, *cue*, encodes the membrane protein CopA and soluble protein CueO. Both would be found to share the substrate Cu(I), however CopA is a P-type ATPase, resting in the inner membrane, while CueO is a periplasmic multi-copper oxidase. CopA and CueO are both

regulated by CueR, which is a transcriptional activator induced by copper.[17] The role of *cue* and the proteins it encodes provide key insights to the cooperative nature of copper resistance in *E. coli*.[28,64,74,76,85,88,92,93,95,113,114,129]

At first, the CusCFBA efflux system was not believed to be related to copper-resistance at all. Franke et al. were not able to distinguish the sensitivity of Δ*cusA* to $CuCl_2$ from that of the wild-type, under aerobic conditions.[25] Grass and Rensing also showed that the MICs of $CuCl_2$ in Δ*cusA* and wild-type strains were indistinguishable (3.5 mM) and, additionally, that there was no observed difference between the growth of Δ*cusA::cm* and the double mutant Δ*cusA::cmcopA::km* in the copper medium.[28] However, their susceptibility assays indicated a slight MIC difference of $CuCl_2$ in *copA::km*Δ*cueO::cm* and *copA::km* (2.5 mM compared to 2.25 mM, respectively), a large difference in Δ*cueO::cm* and Δ*cusA::cm*Δ*cueO::cm* (2.75 mM compared to 1.3 mM, respectively) and no difference at all between Δ*cueO*Δ*cusA::cm*, *copA::km*Δ*cueOcusA::cm* and *copA::km*Δ*cueOcusCFBA::cm*.

It is clear from these results that CueO and particularly CopA each have a major role in conferring aerobic copper resistance in *E. coli,* whereas CusA alone had little to no role. Either the expression of CueO and CopA should be much higher at the copper concentrations tested or they should be much more efficient transporters. Indeed, in *P. syringae*, CopA was shown to bind as many as 11 copper atoms.[15] In addition, *E. coli* CueO was found to have at least five copper binding sites per molecule, with K_m nearly 4-fold higher and k_{cat} 8- and 30-fold higher, respectively, than for yeast Fet3 and human ceruloplasmin.[113] Expression levels seem to fit this scheme, as well. Thieme et al. found that after addition of 1mM $CuCl_2$ expression reached a maximum of 334, 180 and 8 transcripts per cell, of CopA, CueO and CusA, respectively.[122]

Although CusA expression had no noticeable effect on aerobic copper resistance compared to the wild-type cell, a striking difference was observed between the aerobic and anaerobic growth of *E. coli* knockout strain *ΔcusR*.[85] In the absence of oxygen, the *ΔcusR* strain was remarkably sensitive to Cu(I); more than the wild-type and *ΔcueO*, and nearly identical to *ΔcopA*. Under aerobic conditions, the half-maximal inductions of both *cueO* and *copA* were found to occur at 3 μM $CuSO_4$ and *cusC* at 200 μM.[85] In stark contrast to the aerobic case, the half-maximal anaerobic inductions of each protein were found to be 60, 70 and 3 μM for *cueO*, *cusC* and *copA*, respectively. While the expression level of CopA was unchanged, CueO expression fell and CusC expression increased dramatically in the absence of oxygen.

In an oxidizing environment, the dominant copper species should be Cu(II) and in the reducing environment it should be Cu(I). However, because Ag(I) is the dominant silver ion species with or without oxygen, there should be no difference in the aerobic or anaerobic silver-resistance conferred by CusA. The susceptibility experiments of Franke et al. were among the first to confirm the importance of CusA in Ag(I) resistance. They observed that while the wild-type *E. coli* strain grew in the presence of up to 25 μM Ag(I), the *ΔcusA* knockout strain did not.[26] Additionally, *ΔcusA, ΔcusAΔcopA* and *ΔcusCFBA* knockout strains were very sensitive to $AgNO_3$, compared to Δ*copA* knockouts and the wild-type strain, agreeing with the experiments of Gupta et al.[25,33]

Surprisingly, neither CopA nor CueO were implicated in silver resistance. While copper resistance in *E. coli* can be understood in terms of three synergistic components, regarding silver tolerance CusCFBA has no peers. In fact, the *cus* domain was even

prematurely named *agr* before its role in copper-resistance was known.[110] The presence of periplasmic Ag(I) actually inhibits the activity of the copper transporter CueO.[114]

To determine the importance of each component of the tetrapartite efflux system in conferring copper resistance, various knockout strains of *E. coli* were prepared and grown in a medium containing $CuCl_2$. The MIC was determined for each strain and the results were as follows: *ΔcueO* grew uninhibited in up to 3.25 mM $CuCl_2$, *ΔcueOΔcusA* in 1.5, *ΔcueOΔcusB* in 1.5, *ΔcueOΔcusC* in 1.75, *ΔcueOΔcusCΔtolC* in 1.75 and, finally, *ΔcueOΔcusF* in 2.25.[26] Thus, all of the proteins in the *cus* operon were found to be necessary in conferring the maximal copper resistance *in vivo.* The small effect of CusC deletion was initially proposed to be due to TolC replacement, as in the Acr tripartite efflux system, however the *ΔcueOΔcusCΔtolC* knockout strain was no less resistant to copper. Therefore, CusC should be specific to the CusCFBA pump and not replaceable—at least not by TolC.

That copper resistance was strong no matter which subset of components were expressed, and that full copper resistance could not be conferred without expressing the entire pump, suggests that each piece has an active role in heavy-metal efflux. In the following sections, we will examine these roles, describing the structure and proposed function of each component, piece-by-piece.

CusA—Inner-Membrane Transporter, "the Pump"

Prior to the successful X-ray experiments of Long et al. in 2010, diffraction-quality crystals had been obtained for only two other RND pumps: *E. coli* AcrB[61] and *P. aeruginosa* MexB.[106] Based in part on this experiment, CusA was revealed to be the main substrate-binding transporter of this efflux system.

The apo-crystal structure of CusA was first determined to a resolution of 3.52 Å with 98% of the 1,047 residues (residues 5-504 and 516-1040) included in the final model (pdb: 3K07), and it suggests that CusA exists as a homotrimer (Fig. 3).[62] The tertiary structure contains 12-transmembrane α-helices (TMs), typical of membrane transport proteins.[36] TMs 4, 5, 10 and 11 are confined to the cytoplasm and 2 and 8 to the periplasm, while the rest are embedded within the inner-membrane. Like AcrB and MexB before it, CusA can be divided into six sub-domains: PN1, PN2, PC1 and PC2 which comprise the pore domain and DN and DC which are presumed to form the CusB docking domain. Perhaps the most striking feature of CusA is one α-helix that extends horizontally into the bottom of a cleft created by PC1 and PC2 of the periplasmic domain, roughly dividing the transmembrane and periplasmic domains. Sitting atop this helical divider are three adjacent methionine residues (M573, M623 and M672). These same methionines were identified earlier as three of nine total conserved methionines among similar putative drug transporters, including SilA.[93] Although the exact role of these

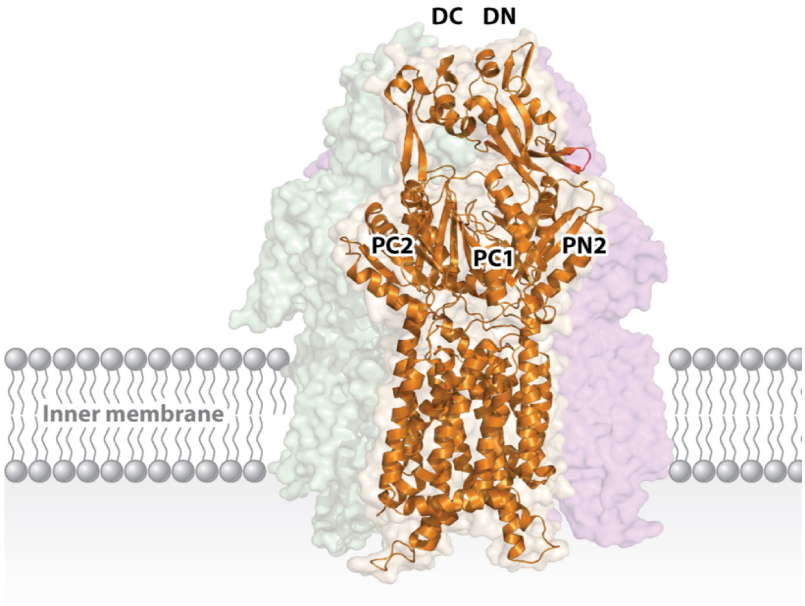

Figure 3. *Crystal structure of the apo-CusA efflux pump. The ribbon diagram (orange) depicts the CusA protomer found in the asymmetric unit of the crystal lattice. The electron density map corresponds to the CusA homotrimer, which is formed by symmetry within the crystal. Subdomains PN1, PN2, PC1, and PC2 form the pore domain, whereas subdomains DN and DC comprise the docking domain, which presumably interacts with the CusC channel.*

methionines was not yet clear, they were found to be indispensable with respect to copper resistance. When mutated to isoleucines, the M573I, M623I and M672I mutants were entirely unable to relieve the copper sensitivity of the *ΔcueO* knockout strain *in vivo*.[93] In the *E. coli ΔacrBΔcueO* knockout strain, Long et al. found that the MICs for $CuSO_4$ in mutants M573I, M623I and M672I were drastically lower than in the wild-type (0.50 mM and 2.25 mM, respectively), while there was no noticeable difference in the MIC level of these mutants when compared to the complete knockout *ΔcusA*. For silver ions, the same effect was also observed. The MIC of $AgNO_3$ in *E. coli* expressing CusA mutants M573I, M623I and M672I was 12.5 μM, compared to 30 μM in the wild-type and 10 μM in the *ΔcusA* knockout strain. In fact, the three-methionine motif is especially common among Cu(I)-resistance proteins,[45,95,127] which suggests the importance of M573, M623 and M672 to the specificity of the CusA pump.

An open question for the almost eight years between experiments was answered when the ion-bound structures of CusA were obtained, by soaking the apo-crystals in Cu(I) and Ag(I) solutions, individually (pdbs: 3KSS and 3KSO, respectively).[62] The root mean square deviation (RMSD) between the Cu(I)- and Ag(I)-bound structures was 1 Å, suggesting that these two structures are nearly identical. Accordingly, Cu(I) and Ag(I) were found to coordinate M573, M623 and M672 in each structure (Fig. 4). In addition to the Cu(I)- and Ag(I)-binding sites, they revealed that metal binding triggers a large conformational change in the pump. The RMSD between the ligand-bound and apo-CusA structures was 3.9 Å. This deviation is mainly attributed to a 30° opening between sub-domains PC1 and PC2, creating

a kind of doorway for metal ions to enter the periplasmic domain of the pump (Fig. 4). The cleft was found closed in the apo-crystal and open in the ligand-bound crystal, suggesting that the binding site is revealed in the presence of Ag(I) or Cu(I) and hidden in their absence. As well as exposing it to the periplasm, the shift more closely coordinates the three methionines involved in binding. In comparison with the apo and ion-bound CusA structures, it is found that transmembrane helices TM1, TM2, TM3, TM6 and TM8 and pore domain PN1 also move noticeably toward the direction of the outer membrane upon metal binding.

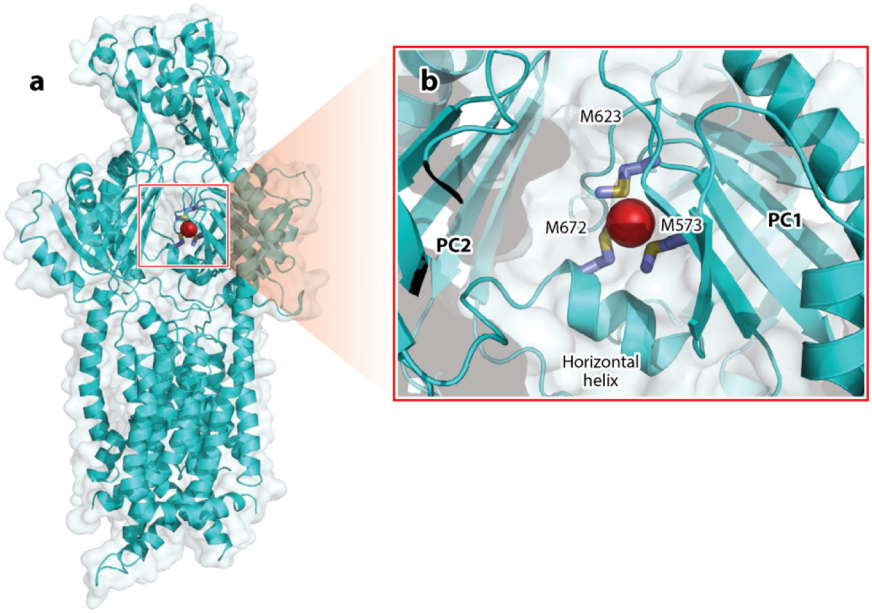

Figure 4. *Crystal structure of the metal ion-bound CusA efflux pump. (a) The ribbon diagram and electron density map correspond to the Cu^+- or Ag^+-CusA subunit (blue). The binding of Cu^+ or Ag^+ by CusA is correlated with a major conformational change in the pump. The separation between the PC1 and PC2 subdomains of each Cu^+- and Ag^+-CusA subunit is exaggerated compared to that of apo-CusA. (b) In the metal ion binding site of CusA, the bound Ag^+ and Cu^+ ions (red) are closely coordinated by methionine residues M623, M672, and M573.*

Besides the main three-methionine binding site, four distinct methionine pairs were identified from each protomer of CusA (Fig. 5). Prior to this experiment, two-methionine or two-cysteine binding sites for copper were identified in at least four other copper tolerance proteins: CusF,[58] CueR,[17] Atx1[5] and CopC.[6,133] In CusA, three of these pairs, M410-M501, M403-M486 and M391-M1009, are found below the main binding site in the transmembrane domain and one, M271-M755, is positioned in the periplasm above the three-methionine site. All of these methionine coordination sites were found on the inside of the channel formed by each protomer of the CusA pump. Based on their locations, it was proposed that the methionine ladder created by these pairs could shuttle ions in a stepwise fashion for ion extrusion. Thus, the periplasmic methionines may serve as a transition of the transported metal ion from CusA to the putative outer-membrane channel CusC.

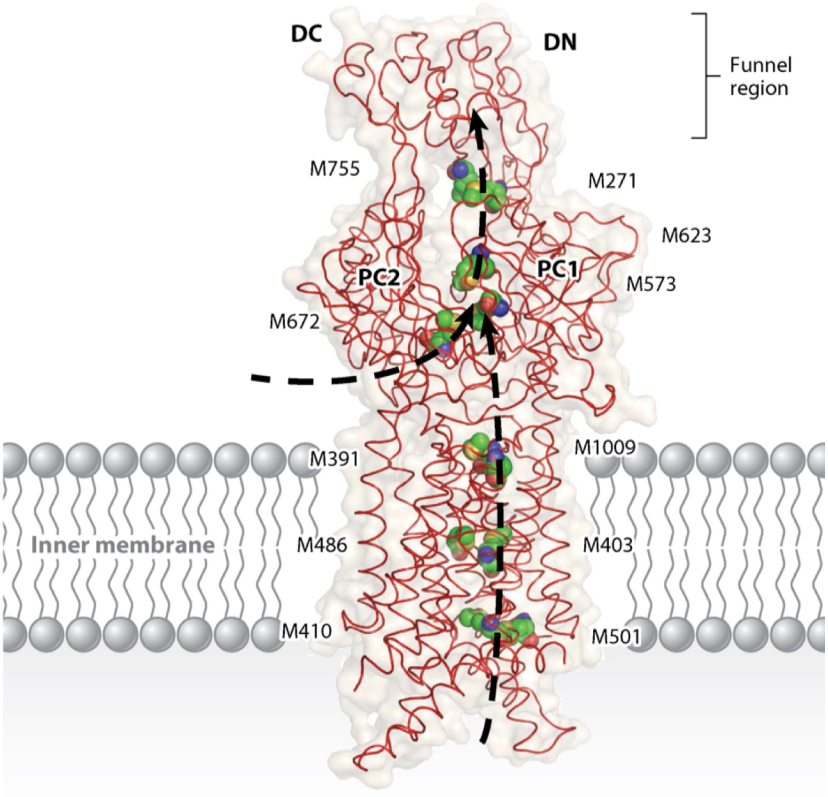

Figure 5. *Metal ion extrusion pathways of the CusA efflux pump. Four methionine pairs depicted by space-filling models (Carbon, green; Oxygen, orange; Nitrogen, blue; Sulfur, yellow) in each protomer of CusA form the path for Cu^{+} or Ag^{+} transport. Metal ions may enter from the cytoplasmic side of the channel via M410-M501 or through the periplasmic cleft between the subdomains PC1 and PC2 via the three-methionine binding sites M573, M623, and M672 (dashed black arrows).*

Susceptibility assays were carried out to determine the utility of these ladder methionines.[62] When one methionine in each pair was mutated to an isoleucine, the MIC of $CuSO_4$ dropped to 1.25 mM for M391I and 1.75 mM for M410I, M486I and M755I, compared to 2.25 mM in the wild-type CusA. Silver resistance was also stunted, as the ladder mutants brought the MIC for $AgNO_3$ to 10 μM, 17.5 μM, 17.5 μM and 12.5 μM for M391I, M410I, M486I and M755I, respectively.

In addition to the *in vivo* susceptibility assays, which demonstrated the importance of each residue in conferring copper and silver resistance, *in vitro* transport assays were employed to determine their roles in metal ion transport. CusA was reconstituted into liposomes containing Phen Green SK and the fluorescent signal was measured in the presence of extravesicular Ag^+. In the wild-type CusA-loaded liposomes, quenching of the fluorescent signal was detected as soon as the silver ions were introduced into the medium. However, in the case of three-methionine binding site mutants M573I, M623I and M672I, and the methionine ladder mutants M391I, M486I and M755I, no quenching was detected, indicating that the extravesicular Ag^+ could not traverse the membrane. Thus, these methionine residues were also implicated in CusA transport, and CusA was shown to have the ability to transport silver directly from the cytoplasm. It should be noted that a similar transport capability has also been found with CzcA[27] and AcrD.[1]

Based on the crystal structure, the dynamics of the trimeric CusA pump was calculated using the elastic network model.[7,62] The result indeed suggests that CusA functions through three coupled motions in which the periplasmic cleft formed by subdomains PC1 and PC2 alternately open and close, similar to AcrB[23,69,70,70,90,103-105] and MexB.[106]

CusB—Membrane Fusion Protein, "the Adaptor"

The structure of apo-CusB, the first of any HME-RND family MFP, was solved two years prior to that of CusA.[62,120] Before the experiments of Su et al. in 2009,[117] both AcrA[67] and MexA[3,39] had produced well-diffracting crystals. Although the CusCFBA complex has many similarities to AcrAB-TolC and MexAB-OprM, CusB itself shares only 13% identity and 52% similarity with MexA and 16% identity and 54% similarity with AcrA.[120]

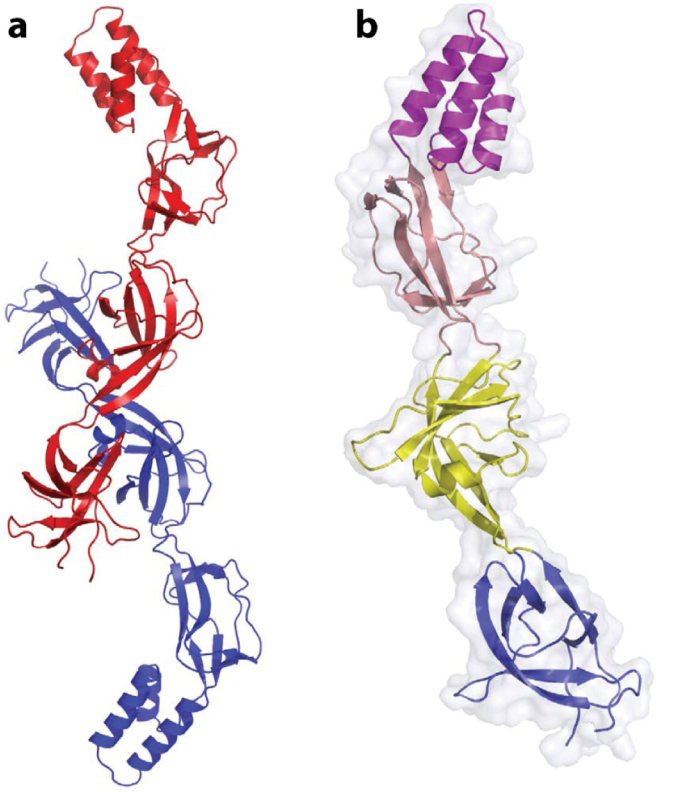

***Figure 6**. Crystal structures of the CusB adaptor. (a) Conformation A (blue) and conformation B (red) of each subunit are depicted by the ribbon diagram. The two distinct structures of the elongated CusB molecule suggest the flexible nature of this protein. (b) Each protomer of CusB can be divided into four domains. An effective hinge between domains 2 and 3 is responsible for the conformational change between conformation A and conformation B of CusB.*

Therefore, we should expect significant differences between these membrane fusion proteins.

The crystal structure of CusB was resolved to 3.4 Å with 78.1% of the total residues (89-385) included (pdb: 3OOC) (Fig. 6).[120] Like AcrA, multiple conformations (here denoted A and B) of CusB were observed inside a single crystal, highlighting the flexibility of this MFP.[67,121] The overall fold, like that of MexA,[3,38,121] can be divided

into four domains: three β-domains and one α-helical domain. The first β-domain was identical in both structures, consisting of the N- and C-terminal ends of the protein arranged in a six β-stranded fashion. This domain was believed to interact directly with the CusA pump. Domain 2 in conformations A and B are quite dissimilar, consisting of two β-strands, four β-sheets and one short α-helix in A and six β-strands and two α-helices in B. Domains 3 and 4 were largely unchanged between each conformation, consisting of eight β-strands and an all α-helical three-helix bundle, respectively. When conformations A and B were superimposed, the overall RMSD calculated between alpha-carbons was 2.6 Å, indicating a significant structural difference between the two molecules. However, the superimposition of domains 1 and 2 of each conformation resulted in an RMSD of only 0.8 Å. Similarly, domains 3 and 4 of A and B could be superimposed to give an RMSD of only 0.8 Å. Therefore, the difference between conformations can be summarized as a shift of domains 1 and 2 with respect to 3 and 4. Best described as a rotation of ~20° about a hinge between the two groups, the difference between conformations results in a shift from what was called the "open" conformation of molecule A to the "closed" conformation of molecule

Several years earlier, in the absence of structural information, Bagai et al. attempted to characterize the binding of Ag(I) and Cu(I) by CusB using isothermal titration calorimetry (ITC), X-ray absorption spectroscopy and *in vitro* susceptibility assays.[8] Titration of Ag(I) into apo-CusB noticeably affected the measured binding enthalpy in a way that was not seen in the control, indicating that CusB does indeed bind Ag(I) *in vitro*. The equilibrium dissociation constant K_d for CusB-Ag(I) binding was measured to be 24.7 nM with a 1:1 Ag(I)-to-CusB molar ratio. Previously, this group had used x-ray absorption spectroscopy to characterize the binding environment of Cu(I) in the CusB-Cu(I) complex and found that the

data was consistent with a three-sulfur coordination. Since there are no cysteine residues in CusB, the bound Cu(I) ion should be coordinated by three proximal methionines, as in CusA.

Of the nine total methionines in CusB, only four were found to be well-conserved and these residues are M21, M36, M38 and M283. Bagai et al. relied on ITC and site-directed mutagenesis to determine if these methionines are important for metal binding. Methionine mutants M21I, M36I, M38I and M283I were prepared, titrated with $Ag(NH_3)_2^+$ and the binding enthalpy was measured. Compared to the wild-type, the mutation of M21 to isoleucine resulted in a 10-fold reduction in binding affinity of CusB for Ag(I). The M36I and M38I mutants resulted in no detectable affinity at all, while M283I had a similar affinity to the wild-type.

This experiment was supplemented with an *in vivo* susceptibility assay using the same CusB mutants, indicating that M21, M36 and M38 probably form a specific binding site for metal ions.[8] Tragically, this three-methionine metal binding site identified by Bagai et al. is located in a region too disordered to obtain meaningful x-ray diffraction data. In combination with their results, the structure of the CusBA adaptor-transporter complex,[119] which we will discuss in the next section, strongly supports the N-terminal residues M21, M36 and M38 as the specific Cu(I)/Ag(I) binding site of CusB.

CusBA—Adaptor-Transporter Complex, "the Co-crystal"

When it was published in 2011, the CusBA co-crystal structure was the first of its kind. This experiment was critical in elucidating specific interactions between the pump and adaptor molecules as well as developing detailed mechanisms of metal ion extrusion by the CusCFBA tetrapartite efflux complex. Its structure was first resolved to 2.90 Å (pdb: 3NE5),

revealing that each CusA monomer interacts with each of two protomers of CusB (Fig. 7).[119] The structure included 1,686 amino acid residues of the total 1,861 (residues 4-1,043 of CusA, residues 79-400 of molecule 1 of CusB and residues 79-402 of molecule 2 of CusB). Based on the co-crystal structure, the trimeric CusA pump is found to contact six CusB

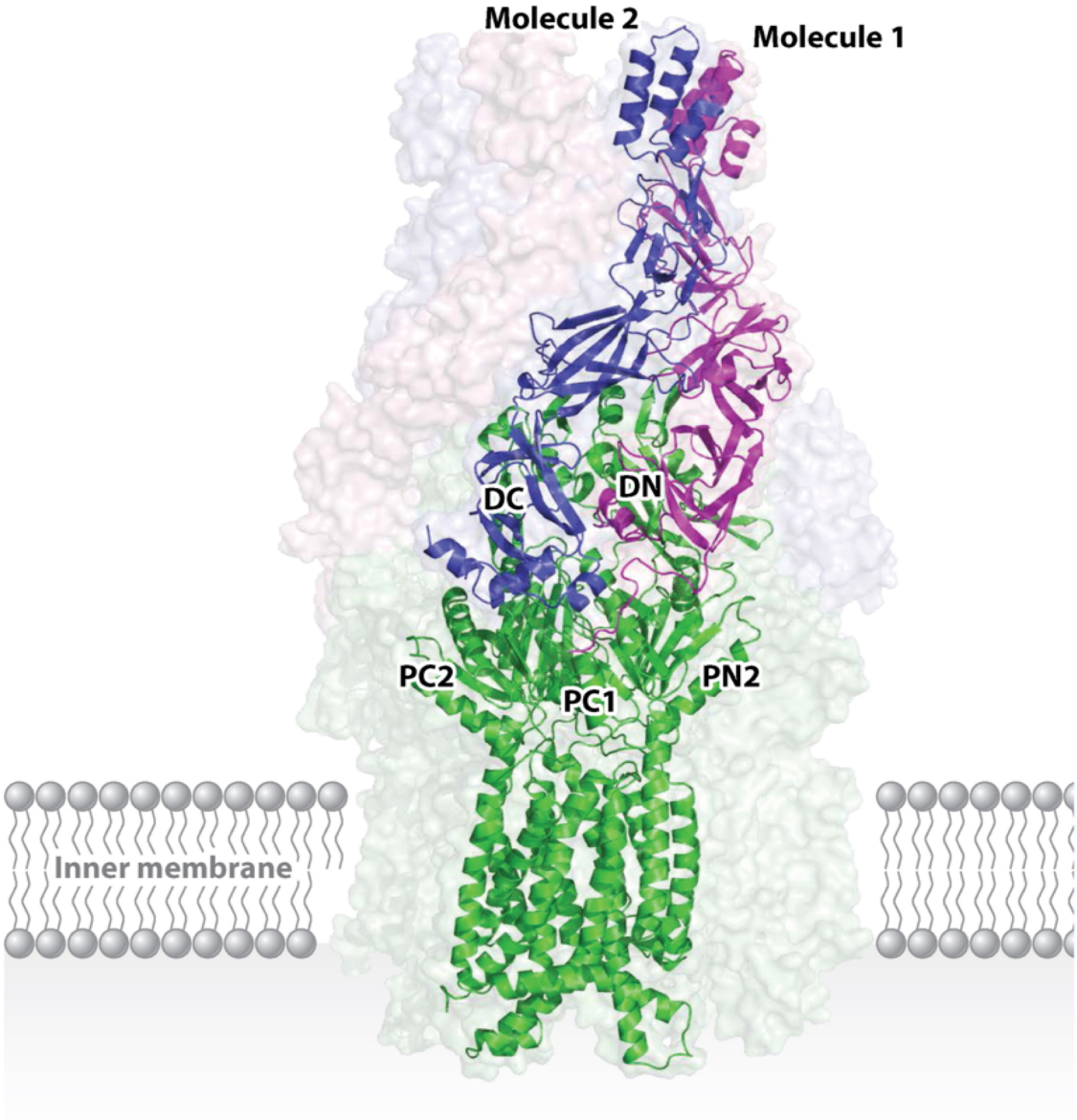

***Figure** 7. Cocrystal structure of the CusBA adaptor-transporter complex. The ribbon diagram depicts the CusBA protomer found in the asymmetric unit of the crystal lattice, and the electron density map corresponds to the CusBA efflux complex. Each subunit of CusBA consists of one CusA molecule (green) and two CusB molecules (magenta and blue). The full structure includes the hexameric CusB adaptor and the trimeric CusA transporter.*

adaptor molecules, which form a hexameric channel right above the periplasmic domain of CusA. The structure of the CusB hexamer mimics an inverted funnel. Domain 1 and the lower half of domain 2 of CusB create a cap-like structure, whereas domains 3, 4, and the upper half of domain 2 form the central channel of the funnel. The inner surface of the cap fits closely on top of the periplasmic domain of the CusA trimer. The funnel-like structure of hexameric CusB creates a channel extending contiguously from the top of the trimeric CusA pump. The narrowest section of the channel is at the hinge of CusB, between domains 2 and 3. The hexameric CusB channel formed above the cap of the adaptor is ~62 Å in length with an average internal diameter of ~37 Å. Thus, the interior of the channel gives rise to a large elongated cavity with a volume of ~65,000 $Å^3$. The lower half of the channel is primarily created by β-barrels, whereas the upper half is an entirely α-helical tunnel. The inner surface of the channel is predominantly negatively charged, as indicated by the electrostatic surface distribution, which suggests that the interior surface of the channel may have the capacity to bind to positively charged metal ions. One striking feature of this possible extrusion pathway is that the inner surface of the hexameric CusB channel is predominantly composed of negatively charged residues, suggesting that the interior surface of the channel may have the capacity to bind to positively charged metal ions.

It is worth noting that the 1:2 relationship of CusA to CusB has led to the conclusion that that the entire assembled structure should be $CusA_3$-$CusB_6$-$CusC_3$. This assembly is indeed in good agreement with the predicted 3:6:3 polypeptide ratios of these three-part complexes.[43,94,115]

ITC was used to determine the binding affinity of CusB to the CusA pump, indicating a tight interaction with a K_d of 5.1 μM. Molecule 1 of CusB interacts predominantly with

CusA through charge–charge interactions, with residues K95, D386, E388 and R397 forming four salt bridges with D155, R771, R777 and E584 of CusA, respectively. Interestingly, the interaction between molecule 2 of CusB and CusA is mediated not by charge-charge but by charge–dipole and dipole–dipole interactions. Specifically, Q108, S109, S253 and N312 of molecule 2 of CusB form hydrogen bonds with Q785, Q194, D800 and Q198 of CusA, respectively.[119]

Recently, artificial peptides, DARPins, have been introduced to bind and effectively inhibit the AcrB transporter.[105] The inhibitor-binding site is found right above the cleft formed between PC1 and PC2 of AcrB. Interestingly, the location of this inhibitor-binding site corresponds to the binding site for molecule 2 of CusB in the CusBA complex. Thus, it is likely that the mechanism of action for these inhibitors may be the disruption of the adaptor-transporter interaction by competition.

By soaking the apo-crystals in Cu(I) solution, three distinct Cu(I)-bound structures were obtained, designated as the pre-extrusion 1, pre-extrusion 2 and extrusion states (pdbs: 3T56, 3T51 and 3T53, respectively).[117] The conformation of the pre-extrusion 1 structure was most similar to the binding state observed in CusA-Cu(I), where the periplasmic cleft formed by PC1 and PC2 was open. In the pre-extrusion 2 state, the conformation of the CusA molecule was more similar to the apo-CusA form (designated as the resting state), with a closed periplasmic cleft and disassembled three-methionine binding site. The extrusion form of CusBA-Cu(I) was nearly identical to the structure of apo-CusBA; similar to the pre-extrusion 2 state, if not for a subtle motion of a short C-terminal helix of molecule 1 of CusB (residues 391-400), which appears to push against another helix of CusA (582-589) and in turn a loop (residues 609-626) directly below the PC1 helix. Based on these interactions, it

was proposed that CusB may be able to tune the width of the channel formed by the CusA pump, optimizing the process of metal ion extrusion.

The conserved charged residues R83, E567, D617, E625, R664 and K678 in the periplasmic domain CusA can be seen in the CusBA structure to form a line along the inner wall of the methionine ladder (Fig. 8). These intriguingly positioned residues were mutated to alanines and expressed in the BL21(DE3) Δ*cueO*Δ*cusA* knockout strain, to determine their importance in copper resistance. The MICs of $CuSO_4$ in *E. coli* carrying these mutant CusA pumps were uniformly 0.5 mM; identical to the empty vector and much less than the 2.25

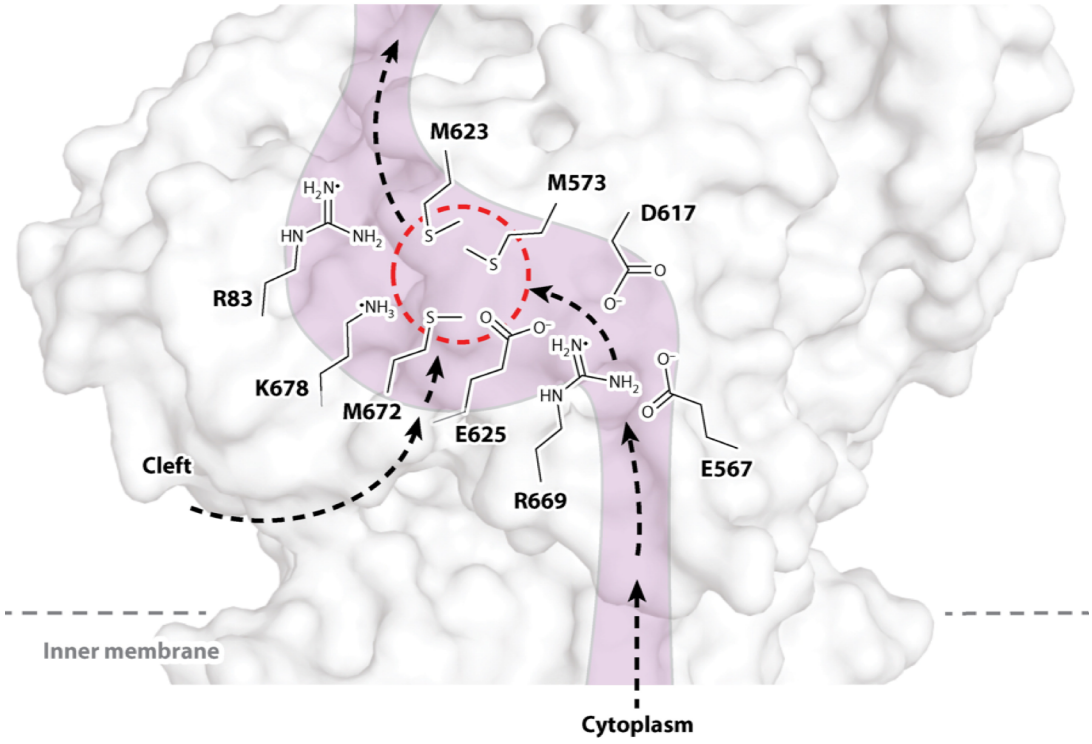

Figure 8. *Cu^+ binding site and conserved charged residues in a schematic representation of the CusA channel. The conserved residues R83, E567, D617, E625, R669, and K678 lining the channel in the periplasmic domain are indicated. The dotted red circle marks the location of the Cu^+ binding site formed by the methionine triad M573, M623, and M672. The paths for metal transport through the periplasmic cleft and transmembrane region are illustrated with curved black arrows.*

mM of the wild-type. An *in vitro* fluorescence assay was also used to determine these residues' importance in silver transport. When liposomes containing the mutant proteins and loaded with fluorescent indicator Phen Green SK were exposed to extravesicular Ag^+ ions no quenching of the signal was observed, suggesting that these mutant transporters were unable to transport Ag^+ across the membrane.

In the earlier work, it was shown that the methionine residues M21, M36 and M38 of CusB represent the most likely Cu(I)- and Ag(I)-binding site. In both crystal structures of the CusB protein and CusBA adaptor-transporter complex, these methionines were located in the disordered N-terminus and their locations could not be resolved. The CusBA co-crystal structure revealed that the disordered N-terminal regions of both CusB molecules are located directly outside the periplasmic cleft of CusA, created by helices PC1 and PC2.[119] Previously, that same cleft was shown to house CusA's single three-methionine binding site.[62] Therefore, it was suggested that CusB may transfer bound Cu(I) and Ag(I) at this location from the periplasm to the CusA pump for extrusion.

CusC—Outer Membrane Factor, "the Channel"

Kulathila et al. published the complete structure (with the exception of disordered residues 21-31) in 2011, to a resolution of 2.3 Å (pdb: 3PIK).[51] Based on the symmetry of the crystal and on gel-filtration measurements, CusC should assemble as a homotrimer, described as "cannon-shaped." Each subunit of CusC in the trimer forms a ~130 Å long α/β barrel, containing four β-strands (contributing to the 12-stranded outer membrane β-barrel) and nine α-helices (forming the elongated periplasmic α-barrel) (Fig. 9). The trimeric CusC channel creates a large cylindrical internal cavity of ~28,000 $Å^3$. Interestingly, the structure

of CusC suggests that the N-terminal cysteine residue of each protomer is covalently linked to the outer membrane via a thioester bond. This N-terminal cysteine residue may play an important role in protein-membrane interactions and could be critical for the insertion of this channel protein into the outer membrane.

The cannon shape is remarkably similar to that of both TolC[50] and OprM,[2] considering that TolC could not replace CusC in the mature pump.[26] Many outer-membrane channels exhibit the character of interchangeability. For example, AcrAB shares TolC with AcrEF and a hemolysin secretion system in *E. coli*,[53] and in *R. metallidurans cnrC* deletion could be functionally restored by *czcC or nccC* expression.[76] However, CusC is believed to only work within the Cus efflux system. This irreplaceability may be partly due to the unique secondary structure

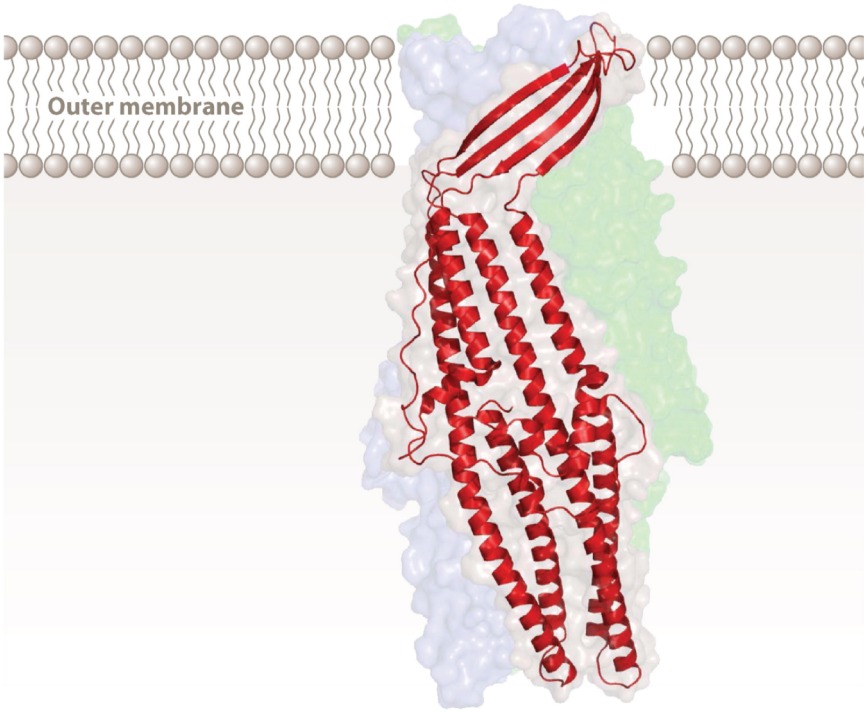

Figure 9. *Crystal structure of the CusC outer membrane transporter. The CusC monomer (red) within the asymmetric unit of the crystal is depicted by the ribbon diagram. The electron density map corresponds to the biological CusC trimer, which is created by crystal symmetry. Each subunit of CusC consists of four β-strands atop nine α-helices, which are arranged as a barrel in the trimeric structure.*

of CusB among members of the MFP family. Specifically, CusB forms a three α-helical domain, which is supposed to directly interact with the CusC channel.

The RMSD calculated over alpha-carbons between CusC and TolC was measured at only 1.6 Å, and between CusC and OprM it was only 1.3 Å. Much of this difference is due to the exceptionally wide extracellular opening of CusC (~30 Å in diameter) compared to those of OprM or TolC (~12 Å). While the extracellular end of each channel is quite different in size, the periplasmic sides are all nearly closed in their respective crystal structures. Therefore, the CusC channel may require a substantial conformational change before allowing extrusion of substrates. Thus far, almost all available crystal structures of outer membrane channels are closed at one or both sides. It would be interesting to visualize the open conformational state of these channel proteins by crystallography.

Although CusC was found to be highly specific to the CusCFBA pump as its outer membrane factor, there was no evidence of copper or silver specificity in the channel. The effects of soaking apo-CusC crystals in high concentrations of Cu(I) or Ag(I), even exceeding 10 mM, could not be observed by X-ray diffraction. Both CusA and CusB were found to coordinate their bound Cu(I) and Ag(I) with methionine residues, however the five methionines in CusC are far removed from one another as seen in the apo-structure. It is unlikely that CusC possesses any binding sites or specificity for Ag(I) and Cu(I) ions, however, as seen in the electrostatic charge distribution, the interior surface of the channel is "strikingly" electronegative.[39] Although not an indication of specificity for metal ions, a similar but less electronegative feature was also observed in the hexameric CusB channel of the crystal structure of the CusBA complex. If the hexameric CusB and trimeric CusC

channels form the extrusion pathway of the pump, the electrostatic gradient created by the two interior surfaces may provide a possible pathway for ion extrusion.

To understand how the CusBA complex interacts with CusC, a CusCBA fitting model was constructed based on the crystal structures of CusBA complex and CusC protein. The final $CusC_3$-$CusB_6$-$CusA_3$ structural model represents an assembled tripartite efflux complex spanning both the inner and outer membranes of *E. coli* to extrude metal ions (Fig. 10).

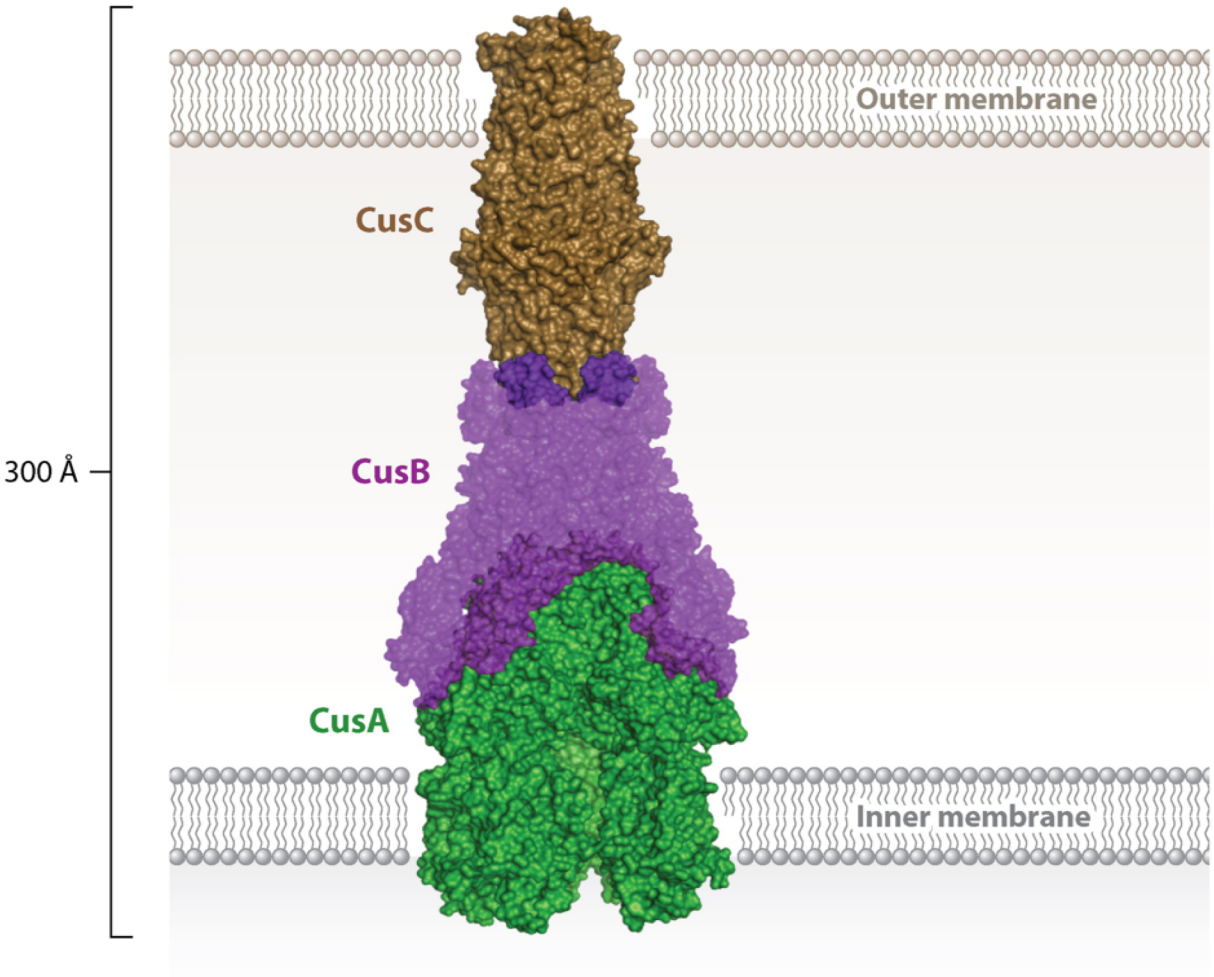

Figure 10. *Docking of CusC to CusBA. The α-helical end of CusC interacts with the α-helices (domain 4) of CusB in the CusBA complex. The surface rendering of the $CusC_3$-$CusB_6$-$CusA_3$ complex is colored as follows: brown, CusC trimer; purple, CusB hexamer; green, CusA trimer.*

CusF—Small Periplasmic Metal-Binding Protein, "the Chaperone"

So far, each component (CusA, CusB and CusC) of our tetrapartite complex has a counterpart in the protypical multi-drug resistance systems AcrAB-TolC and MexAB-OprM. Then, there is CusF. When the *cusF* gene was first sequenced, nothing like it had been found in any known RND systems or any drug transporters.[110] However, by the time its crystal structure was published, more than 35 CusF homologues had been identified among Gram-negative bacteria, including two putative heavy-metal resistance determinants: *czc* of *R. metallidurans* and *sil* of *S. typhimurium*.[59] The function of these proteins was largely unknown; nonetheless, three conserved residues M47, M49 and H36 were identified among their sequences, suggesting that these residues may represent a potential binding site.

The crystal structure of apo-CusF was published by Loftin et al. in 2005 (pdb: 1ZEQ),[59] and two years later the CusF-Ag(I) complex structure was solved by the same group (pdb: 2QCP).[58] The apo-structure was resolved to 1.5 Å, including ~93% of the total residues (residues 6-88). Based on the diffraction results, the structure of CusF can be simply described as a small five-stranded β-barrel, composed of two anti-parallel three-stranded β-sheets which were packed orthogonally. The CusF-Ag(I) complex was resolved to 1.0 Å, including residues 10-88.[58] An RMSD of only 0.5 Å was calculated between the apo and ion-bound structures, indicating that these two structures are essentially the same. The bound Ag(I) in CusF was found to be indeed coordinated by H36, M47 and M49.

The crystal structure of the CusF-Cu(I) complex (pdb: 2VB2) (Fig. 11) was published by Xue et al., one year later, with a resolution of 1.7 Å and including residues 13-87.[128] Their structure was nearly identical to both the previous CusF-Ag(I) complex and apo-CusF structures. However, there was a slight conformational change observed upon Cu(I) binding,

which includes a twist of residues M47 and M49 to accommodate the bound metal ion. Upon ion-binding, these side chains swiveled to more closely coordinate the bound Cu(I), with a fourth residue, W44, implicated in binding. The proximity of this tryptophan residue to Cu(I) (~2.7 Å as determined by x-ray absorption spectroscopy) was "unprecedented" among previous metal-binding sites, and introduces a novel motif for Cu(I)-binding.

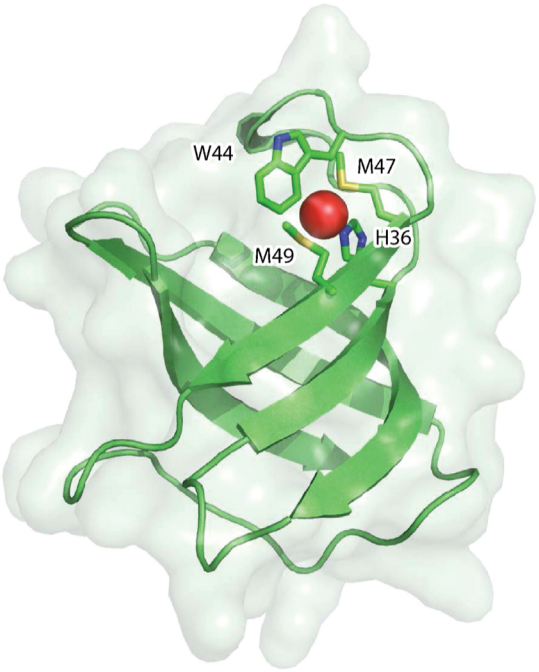

Figure 11. *Crystal structure of the metal ion--bound CusF metallochaperone. The binding of Cu^{+} and Ag^{+} (red) is coordinated by M47, M49, W44, and H36, which form a novel metal ion-binding site in CusF.*

Although it was found to bind both Cu(I) and Ag(I), it was yet unclear what role CusF might have in the efflux process. Recent experiments strongly suggest the possibility of a hand-off of bound Cu(I) or Ag(I) from CusF to CusB.[9,16,66] Presumably, the specific interaction between CusF and CusB is critical for a direct transfer of Cu(I) and Ag(I) ions before extrusion via the CusCFBA efflux complex. Thus, the interaction between these two components may mark the initial step for metal ion extrusion by the CusCFBA tetrapartite efflux complex.

Concluding Remarks

It has been well established that overexpression of RND multi-drug efflux pumps led to a resistant phenotype in pathogenic organisms. This problem is exacerbated by the ease with which many of these resistance genes can pass from one microorganism to another through plasmid transfer.[20] Because of the fact that these multi-drug efflux pumps are able to respond to a wide spectrum of substrates, pathogenic bacteria that overexpress these pumps can be selected for by many different agents. For example, the organic compound nonoxynol-9, which is widely used as an active ingredient in most spermicide creams and jellies, can select *Neisseria gonorrhoeae* mutants that overexpress the MtrCDE RND-type multi-drug efflux pump.[125] Thus, it is very important to understand the molecular mechanism as well as detailed structural information of these efflux pumps to combat infectious diseases.

The availability of the three-dimensional structures of these efflux transporters potentially allows us to rationally design agents that block their function. However, there is still quite a mountain to climb in the development toward achieving this goal. For instance, a crystallographic model of these RND tripartite efflux complexes has not yet been reported. One direct approach to obtain the complete picture of these efflux complexes is to elucidate the structures of individual components as well as the assembled complexes of these efflux systems. This invaluable structural information should be able to provide a platform for producing new drugs and inhibitors to heighten bacterial sensitivity to these antimicrobials. In addition, the outer membrane channel proteins of these RND efflux complexes that span the outer membrane and expose to the surface of bacterial cells make them ideal targets for vaccines. Thus, the knowledge of these outer membrane channels is essential for the development of vaccines against bacterial diseases.

To date, no efflux pump inhibitor has been licensed for use in the clinical treatment of bacterial infections. Bacterial efflux pump inhibitors may probably simultaneously inhibit human detoxification systems and bacterial multi-drug efflux pumps. This will require a thorough research to ensure that these inhibitors are specific solely for the bacterial efflux pumps, minimizing their side effects to the host. As RND pumps assemble as tripartite complexes, one method to inhibit the drug-resistant phenotype is to prevent the assembly of complex formation by blocking the different components from forming a functional complex. The possibility of this approach has been demonstrated by the crystal structure of AcrB in complex with the designed ankyrin repeat protein inhibitor (DARPin).[105] Based on the structural information of the CusBA adaptor-transporter complex,[119] the inhibitor in principle is bound in such a way that the AcrB transporter is not able to form a functional complex with the AcrA adaptor. Therefore, the structural information of multi-drug efflux pumps, particularly with their assembled complex models, is of importance. In the case of the Cus efflux system, the availability of the co-crystal structures of the adaptor-transporter CusBA, channel-adaptor CusCB, and complete tetrapartite CusCFBA efflux complexes are essential in order to achieve this task although co-crystallization of different components of these RND efflux systems have been proven to be extremely difficult.

Another potential approach is to target the transcriptional regulators that modulate the expression of these RND pumps. For example, the global regulator CmeR controls the expression of several genes, including *cmeABC* which produces the CmeABC RND efflux complex,[30,55-57] and the transcriptional repressor MtrR which regulates the expression level of the MtrCDE tripartite multi-drug efflux pump.[35,40,107,125] By inhibiting the interactions between the regulators and their corresponding substrates, expression of these transporter

genes could potentially be blocked. Recently, more examples of RND transporters regulated by two-component systems have been identified, including *E. coli* CusRS, which controls the expression of the CusCFBA efflux complex.[31,68] The structures of different components of these regulatory systems may potentially allow us to rationally design inhibitors that reduce the level of efflux pump expression.

References

1. Aires JR, Nikaido H. 2005. Aminoglycosides are captured from both periplasm and cytoplasm by the AcrD multidrug efflux transporter of *Escherichia coli*. *J. Bacteriol.* 187(6):1923—29

2. Akama H, Kanemaki M, Yoshimura M, Tsukihara T, Kashiwagi T, et al. 2004. Crystal structure of the drug discharge outer membrane protein, OprM, of *Pseudomonas aeruginosa*: dual modes of membrane anchoring and occluded cavity end. *J. Biol. Chem.* 279(51):52816—19

3. Akama H, Matsuura T, Kashiwagi S, Yoneyama H, Narita S, et al. 2004. Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in *Pseudomonas aeruginosa*. *J. Biol. Chem.* 279(25):25939—42

4. Angus BL, Carey AM, Caron DA, Kropinski AM, Hancock RE. 1982. Outer membrane permeability in *Pseudomonas aeruginosa*: comparison of a wild-type with an antibiotic-supersusceptible mutant. *Antimicrob. Agents Chemother.* 21(2):299—309

5. Arnesano F, Banci L, Bertini I, Huffman DL, O'Halloran TV. 2001. Solution structure of the Cu(I) and apo forms of the yeast metallochaperone, Atx1. *Biochemistry* 40(6):1528--39

6. Arnesano F, Banci L, Bertini I, Mangani S, Thompsett AR. 2003. A redox switch in CopC: an intriguing copper trafficking protein that binds copper(I) and copper(II) at different sites. *Proc. Natl. Acad. Sci. USA* 100(7):3814--19

7. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, et al. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80(1):505--15

8. Bagai I, Liu W, Rensing C, Blackburn NJ, McEvoy MM. 2007. Substrate-linked conformational change in the periplasmic component of a Cu(I)/Ag(I) efflux system. *J. Biol. Chem.* 282(49):35695--702

9. Bagai I, Rensing C, Blackburn NJ, McEvoy MM. 2008. Direct metal transfer between periplasmic proteins identifies a bacterial copper chaperone. *Biochemistry* 47(44):11408--14

10. Baranova N, Nikaido H. 2002. The *baeSR* two-component regulatory system activates transcription of the *yegMNOB* (*mdtABCD*) transporter gene cluster in *Escherichia coli* and increases its resistance to novobiocin and deoxycholate. *J. Bacteriol.* 184(15):4168--76

11. Blattner FR, Plunkett G III, Bloch CA, Perna NT, Burland V, et al. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453--62

12. Bohnert JA, Schuster S, Fähnrich E, Trittler R, Kern WV. 2007. Altered spectrum of multidrug resistance associated with a single point mutation in the *Escherichia coli* RND-type MDR efflux pump YhiV (MdtF). *J. Antimicrob. Chemother.* 59(6):1216--22

13. Brown MH, Paulsen IT, Skurray RA. 1999. The multidrug efflux protein NorM is a prototype of a new family of transporters. *Mol. Microbiol.* 31:394--95

14. Brown NL, Rouch DA, Lee BTO. 1992. Copper resistance determinants in bacteria. *Plasmid* 27:41--51

15. Cha J-S, Cooksey DA. 1991. Copper resistance in *Pseudomonas syringae* mediated by periplasmic and outer membrane proteins. *Proc. Natl. Acad. Sci. USA* 88:8915--19

16. Chakravorty DK, Wang B, Ucisik MN, Merz KM Jr. 2011. Insight into the cation-π interaction at the metal binding site of the copper metallochaperone. *J. Am. Chem. Soc.* 133(48):199330--33

17. Changela A, Chen K, Xue Y, Holschen J, Outten CE, et al. 2003. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* 301(5638):1383--87

18. Chopra I. 2007. The increasing use of silver-based products as antimicrobial agents: a useful development or a cause for concern? *J. Antimicrob. Chemother.* 59:587--90

19. Conroy O, Kim E-H, McEvoy MM, Rensing C. 2010. Differing ability to transport non-metal substrates by two RND-type metal exporters. *FEMS Microbiol. Lett.* 308(2):115--22

20. Courvalin P. 1994. Transfer of antibiotic resistance genes between gram-positive and gram-negative bacteria. *Antimicrob. Agents Chemother.* 38:1447--51

21. Das D, Xu QS, Lee JY, Ankoudinova I, Huang C, et al. 2007. Crystal structure of the multidrug efflux transporter AcrB at 3.1 Å resolution reveals the N-terminal region with conserved amino acids. *J. Struct. Biol.* 158(3):494--502

22. Delmar JA, Su C-C, Yu EW. 2013. Structural mechanisms of heavy-metal extrusion by the Cus efflux system. *Biometals* 26(4):593--607

23. Eicher T, Cha HJ, Seeger MA, Brandstätter L, El-Delik J, et al. 2012. Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proc. Natl. Acad. Sci. USA* 109(15):5687--92

24. Festa RA, Thiele DJ. 2011. Copper: an essential metal in biology. *Curr. Biol.* 21(21):R877--83

25. Franke S, Grass G, Nies DH. 2001. The product of the *ybdE* gene of the *Escherichia coli* chromosome is involved in detoxification of silver ions. *Microbiology* 147:965--72

26. Franke S, Grass G, Rensing C, Nies DH. 2003. Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *J. Bacteriol.* 185(13):3804--12

27. Goldberg M, Pribyl T, Juhnke S, Nies DH. 1999. Energetics and topology of CzcA, a cation/proton antiporter of the resistance-nodulation-cell division protein family. *J. Biol. Chem.* 273(37):26065--70

28. Grass G, Rensing C. 2001. Genes involved in copper homeostasis in *Escherichia coli*. *J. Bacteriol.* 183(6):2145--47

29. Grass G, Rensing C, Solioz M. 2011. Metallic copper as an antimicrobial surface. *Appl. Environ. Microbiol.* 77(5):1541--47

30. Gu R, Su C-C, Shi F, Li M, McDermott G, et al. 2007. Crystal structure of the transcriptional regulator CmeR from *Campylobacter jejuni*. *J. Mol. Biol.* 372(3):583--93

31. Gudipaty SA, Larsen AS, Rensing C, McEvoy MM. 2012. Regulation of Cu(I)/Ag(I) efflux genes in *Escherichia coli* by the sensor kinase CusS. *FEMS Microbiol. Lett.* 330:30--37

32. Gupta A, Matsui K, Lo J-F, Silver S. 1999. Molecular basis for resistance to silver cations in *Salmonella*. *Nat. Med.* 5(2):183--88

33. Gupta A, Phung LT, Taylor DE, Silver S. 2001. Diversity of silver resistance genes in IncH incompatibility group plasmids. *Microbiology* 147:3393--402

34. Haber F, Weiss J. 1932. On the catalysis of hydroperoxides. *Naturwissenschaften* 20(51):948--50

35. Hagman KE, Shafer WM. 1995. Transcriptional control of the *mtr* efflux system of *Neisseria gonorrhoeae*. *J. Bacteriol.* 177(14):4162--65

36. Henderson PJ. 1993. The 12-transmembrane helix transporters. *Curr. Opin. Cell Biol.* 5(4):708--21

37. Higgins CF. 2007. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* 446:749--57

38. Higgins MK, Bokma E, Koronakis E, Hughes C, Koronakis V. 2004. Structure of the periplasmic component of a bacterial drug efflux pump. *Proc. Natl. Acad. Sci. USA* 101(27):9995--99

39. Hinchliffe P, Symmons MF, Hughes C, Koronakis V. 2013. Structure and operation of bacterial tripartite pumps. *Annu. Rev. Microbiol.* 67:221--42

40. Hoffmann KM, Williams D, Shafer WM, Brennan RG. 2005. Characterization of the multiple transferable resistance repressor MtrR, from *Neisseria gonorrhoea*. *J. Bacteriol.* 187(14):5008--12

41. Hooper DC. 1999. Mechanisms of fluoroquinolone resistance. *Drug Resist. Update* 2:38--55

42. Jaffé A, Chabbert YA, Derlot E. 1983. Selection and characterization of β-lactam-resistant *Escherichia coli* K-12 mutants. *Antimicrob. Agents Chemother.* 23(4):622--25

43. Janganan TK, Bavro VN, Zhang L, Matak-Vinkovic D, Barrera NP, et al. 2011. Evidence for the assembly of a bacterial tripartite multidrug pump with a stoichiometry of 3:6:3. *J. Biol. Chem.* 286(30):26900--12

44. Janganan TK, Zhang L, Bavro VN, Matak-Vinkovic D, Barrera NP, et al. 2011. Opening of the outer membrane protein channel in tripartite efflux pumps is induced by interaction with the membrane fusion partner. *J. Biol. Chem.* 286(7):5484--93

45. Jiang J, Nadas IA, Kim MA, Franz KJ. 2005. A Mets motif peptide found in copper transport proteins selectively binds Cu(I) with methionine-only coordination. *Inorg. Chem.* 44(26):9787--94

46. Kim E-H, Nies DH, McEvoy MM, Rensing C. 2011. Switch or funnel: how RND-type transport systems control periplasmic metal homeostasis. *J. Bacteriol.* 193(10):2381--87

47. Kim HS, Nikaido H. 2012. Different function of MdtB and MdtC subunits in the heterotrimeric efflux transporter $MdtB_2C$ complex of *Escherichia coli*. *Biochemistry* 51(20):4188--97

48. Kittleson JT, Loftin IR, Hausrath AC, Engelhardt KP, Rensing C, McEvoy MM. 2006. Periplasmic metal-resistance protein CusF exhibits high affinity and specificity for both Cu^I and Ag^I. *Biochemistry* 45(37):11096--102

49. Kobayashi N, Nishino K, Yamaguchi A. 2001. Novel macrolide-specific ABC-type efflux transporter in *Escherichia coli*. *J. Bacteriol.* 183(19):5639--44

50. Koronakis V, Sharff A, Koronakis E, Luisi B, Hughes C. 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 405:914--19

51. Kulathila R, Kulathila R, Indic M, van den Berg B. 2011. Crystal structure of *Escherichia coli* CusC, the outer membrane component of a heavy metal efflux pump. *PLoS ONE* 6(1):e15610

52. Lane TW, Saito MA, George GN, Pickering IJ, Prince RC, et al. 2005. Biochemistry: a cadmium enzyme from a marine diatom. *Nature* 435:42

53. Lau SY, Zgurskaya HI. 2005. Cell division defects in *Escherichia coli* deficient in the multidrug efflux transporter AcrEF-TolC. *J. Bacteriol.* 187(22):7815--25

54. Levy SB. 1992. Active efflux mechanisms for antimicrobial resistance. *Antimicrob. Agents Chemother.* 36(4):695--703

55. Li XZ, Nikaido H. 2004. Efflux-mediated drug resistance in bacteria. *Drugs* 64(2):159--204

56. Lin J, Akiba M, Sahin O, Zhang Q. 2005. CmeR functions as a transcriptional repressor for the multidrug efflux pump CmeABC in *Campylobacter jejuni*. *Antimicrob. Agents Chemother.* 49(3):1067--75

57. Lin J, Michel LO, Zhang Q. 2002. CmeABC functions as a multidrug efflux system in *Campylobacter jejuni*. *Antimicrob. Agents Chemother.* 46(7):2124--31

58. Loftin IR, Franke S, Blackburn NJ, McEvoy MM. 2007. Unusual Cu(I)/Ag(I) coordination of *Escherichia coli* CusF as revealed by atomic resolution crystallography and X-ray absorption spectroscopy. *Prot. Sci.* 16:2287--93

59. Loftin IR, Franke S, Roberts SA, Weichsel A, Héroux A, et al. 2005. A novel copper-binding fold for the periplasmic copper resistance protein CusF. *Biochemistry* 44(31):10533--40

60. Loftin IR, McEvoy MM. 2009. Tryptophan Cu(I)-π interaction fine-tunes the metal binding properties of the bacterial metallochaperone CusF. *J. Biol. Inorg. Chem.* 14(6):905--12

61. Long F, Su C-C, Lei H-T, Bolla JR, Do SV, Yu EW. 2012. Structure and mechanism of the tripartite CusCBA heavy-metal efflux complex. *Phil. Trans. R. Soc. B* 367:1047--58

62. Long F, Su C-C, Zimmermann MT, Boyken SE, Rajashankar KR, et al. 2010. Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport. *Nature* 467(7314):484--88

63. Ma D, Cook DN, Alberti M, Pon NG, Nikaido H, et al. 1993. Molecular cloning and characterization of *acrA* and *acrE* genes of *Escherichia coli*. *J. Bacteriol.* 175(19):6299--313

64. Macomber L, Imlay JA. 2009. The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proc. Natl. Acad. Sci. USA* 106(20):8344--49

65. McHugh GL, Moellering RC, Hopkins CC, Swartz MN. 1975. *Salmonella typhimurium* resistant to silver nitrate, chloramphenicol, and ampicillin. *Lancet* 305(7901):235--40

66. Mealman TD, Zhou M, Affandi T, Chacón KN, Aranguren ME, et al. 2012. N-terminal region of CusB is sufficient for metal binding and metal transfer with the metallochaperone CusF. *Biochemistry* 51:6767--75

67. Mikolosko J, Bobyk K, Zgurskaya HI, Ghosh P. 2006. Conformational flexibility in the multidrug efflux system protein AcrA. *Structure* 14(3):577--87

68. Munson GP, Lam DL, Outten FW, O'Halloran TV. 2000. Identification of a copper-responsive two-component system on the chromosome of *Escherichia coli* K-12. *J. Bacteriol.* 182(20):5864--71

69. Murakami S. 2008. Multidrug efflux transporter, AcrB---the pumping mechanism. *Curr. Opin. Struc. Biol.* 18:459--65

70. Murakami S, Nakashima R, Yamashita E, Matsumoto T, Yamaguchi A. 2006. Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* 443:173--79

71. Murakami S, Nakashima R, Yamashita E, Yamaguchi A. 2002. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* 419:587--93

72. Nagakubo S, Nishino K, Hirata T, Yamaguchi A. 2002. The putative response regulator BaeR stimulates multidrug resistance of *Escherichia coli* via a novel multidrug exporter system, MdtABC. *J. Bacteriol.* 184(15):4161--67

73. Nies DH. 1995. The cobalt, zinc, and cadmium efflux system CzcABC from *Alcaligenes eutrophus* functions as a cation-proton antiporter in *Escherichia coli*. *J. Bacteriol.* 177(10):2707--12

74. Nies DH. 1999. Microbial heavy-metal resistance. *Appl. Microbiol. Biotechnol.* 51:730--50

75. Nies DH. 2003. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol. Rev.* 27(2):313--39

76. Nies DH. 2007. Bacterial transition metal homeostasis. In *Molecular Microbiology of Heavy Metals*, vol. 6, ed. DH Nies, S Silver, pp. 117--42. Berlin Heidelberg: Springer

77. Nies DH. 2013. RND efflux pumps for metal cations. In *Microbial Efflux Pumps: Current Research*, Vol. 1, ed. EW Yu, Q Zhang, MH Brown, pp. 79--121. Norfolk, UK: Caister Academic

78. Nies DH, Nies A, Chu L, Silver S. 1989. Expression and nucleotide sequence of a plasmid-determined divalent cation efflux system from *Alcaligenes eutrophus*. *Proc. Natl. Acad. Sci. USA* 86:7351--55

79. Nikaido H. 1989. Outer membrane barrier as a mechanism of antimicrobial resistance. *Antimicrob. Agents Chemother.* 33(11):1831--36

80. Nikaido H. 1994. Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science* 264:382--88

81. Nikaido H. 1996. Multidrug efflux pumps of gram-negative bacteria. *J. Bacteriol.* 178(20):5853--59

82. Nikaido H. 2011. Structure and mechanism of RND-type multidrug efflux pumps. *Adv. Enzymol. Relat. Areas Mol. Biol.* 77:1--60

83. Nikaido H, Takatsuka Y. 2009. Mechanisms of RND multidrug efflux pumps. *Biochim. Biophys. Acta* 1794:769--81

84. Nishino K, Yamaguchi A. 2001. Analysis of a complete library of putative drug transporter genes in *Escherichia coli*. *J. Bacteriol.* 183(20):5803--12

85. Outten FW, Huffman DL, Hale JA, O'Halloran TV. 2001. The independent *cue* and *cus* systems confer copper tolerance during aerobic and anaerobic growth in *Escherichia coli*. *J. Biol. Chem.* 276(33):30670--77

86. Paulsen IT, Brown MH, Skurray RA. 1996. Proton-dependent multi-drug efflux systems. *Microbiol. Rev.* 60(4):575--608

87. Paulsen IT, Nguyen L, Sliwinski MK, Rabus R, Saier MH Jr. 2000. Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *J. Mol. Biol.* 301:75--100

88. Peña MM, Lee J, Thiele DJ. 1999. A delicate balance: homeostatic control of copper uptake and distribution. *J. Nutr.* 129(7):1251--60

89. Percival SL, Bowler PG, Russell D. 2005. Bacterial resistance to silver in wound care. *J. Hosp. Infect.* 60:1--7

90. Pos KM. 2009. Drug transport mechanism of the AcrB efflux pump. *Biochim. Biophys. Acta* 1794(5):782--93

91. Pugsley AP, Schnaitman CA. 1978. Identification of three genes controlling production of new outer membrane pore proteins in *Escherichia coli* K-12. *J. Bacteriol.* 135(3):1118--29

92. Rensing C, Fan B, Sharma R, Mitra B, Rosen BP. 2000. CopA: an *Escherichia coli* Cu(I)-translocating P-type ATPase. *Proc. Natl. Acad. Sci. USA* 97(2):652--56

93. Rensing C, Grass G. 2003. *Escherichia coli* mechanisms of copper homeostasis in a changing environment. *FEMS Microbiol. Rev.* 27:197--213

94. Rensing C, Pribyl T, Nies DH. 1997. New functions for the three subunits of the CzcCBA cation-proton antiporter. *J. Bacteriol.* 179:6871--79

95. Roberts SA, Weichsel A, Grass G, Thakali K, Hazzard JT, et al. 2002. Crystal structure and electron transfer kinetics of CueO, a multicopper oxidase required for copper homeostasis in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 99(5):2766--71

96. Rosenberg EY, Ma D, Nikaido H. 2000. AcrD of *Escherichia coli* is an aminoglycoside efflux pump. *J. Bacteriol.* 182(6):1754--56

97. Routh MD, Zalucki Y, Su C-C, Long F, Zhang Q, et al. 2011. Efflux pumps of the resistance-nodulation-division family: a perspective of their structure, function, and regulation in gram-negative bacteria. In *Advances in Enzymology and Related Areas of Molecular Biology*, Vol. 77, ed. EJ Toone, pp. 109--46. Hoboken, NJ: Wiley

98. Russell AD, Hugo WB. 1994. Antimicrobial activity and action of silver. *Prog. Med. Chem.* 31:351--70

99. Saier MH Jr. 1994. Computer-aided analyses of transport protein sequences: gleaning evidence concerning function, structure, biogenesis, and evolution. *Microbiol. Rev.* 58(1):71—93

100. Saier MH Jr. 1998. Molecular phylogeny as a basis for the classification of transport proteins from bacteria, archaea and eukarya. *Adv. Microb. Physiol.* 40(81):81—136

101. Saier MH Jr, Beatty JT, Goffeau A, Harley KT, Heijne WH, et al. 1999. The major facilitator superfamily. *J. Mol. Microbiol. Biotechnol.* 1(2):257—79

102. Saier MH Jr, Paulsen IT. 2001. Phylogeny of multidrug transporters. *Semin. Cell Dev. Biol.* 12:205—13

103. Seeger MA, Diederichs K, Eicher T, Brandstätter L, Schiefner A, et al. 2008. The AcrB efflux pump: Conformational cycling and peristalsis lead to multidrug resistance. *Curr. Drug Targets* 9(9):729—49

104. Seeger MA, Schiefner A, Eicher T, Verrey F, Diederichs K, et al. 2006. Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science* 313(5791):1295—98

105. Sennhauser G, Amstutz P, Briand C, Storchenegger O, Grütter MG. 2007. Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS ONE* 5(1):e7

106. Sennhauser G, Bukowksa MA, Briand C, Grütter MG. 2009. Crystal structure of the multidrug exporter MexB from *Pseudomonas aeruginosa*. *J. Mol. Biol.* 389:134--45

107. Shafer WM, Qu X-D, Waring AJ, Lehrer RI. 1998. Modulation of the *Neisseria gonorrhoeae* susceptibility to vertebrate antibacterial peptides due to a member of the resistance/nodulation/division efflux pump family. *Proc. Natl. Acad. Sci. USA* 95(4):1829--33

108. Silver S. 1996. Bacterial heavy metal resistance: new surprises. *Annu. Rev. Microbiol.* 50:753--89

109. Silver S. 1996. Bacterial resistances to toxic metal ions---a review. *Gene* 179:9--19

110. Silver S. 2003. Bacterial silver resistance: molecular biology and uses and misuses of silver compounds. *FEMS Microbiol. Rev.* 27:341--53

111. Silver S, Misra TK. 1988. Plasmid-mediated heavy metal resistances. *Annu. Rev. Microbiol.* 42:717--43

112. Silver S, Phung LT, Silver G. 2006. Silver as biocides in burn and wound dressings and bacterial resistance to silver compounds. *J. Ind. Microbiol. Biotechnol.* 33:627--34

113. Singh SK, Grass G, Rensing C, Montfort WR. 2004. Cuprous oxidase activity of CueO from *Escherichia coli*. *J. Bacteriol.* 186(22):7815--17

114. Singh SK, Roberts SA, McDevitt SF, Weichsel A, Wildner GF, et al. 2011. Crystal structures of multicopper oxidase CueO bound to copper(I) and silver(I): functional role of a methionine-rich sequence. *J. Biol. Chem.* 286(43):37849--57

115. Stegmeier JF, Polleichtner G, Brandes N, Hotz C, Andersen C. 2006. Importance of the adaptor (membrane fusion) protein hairpin domain for the functionality of multidrug efflux pumps. *Biochemistry* 45:10303--12

116. Su C-C, Li M, Gu R, Takatsuka Y, McDermott G, et al. 2006. Conformation of the AcrB multidrug efflux pump in mutants of the putative proton relay pathway. *J. Bacteriol.* 188(20):7290--96

117. Su C-C, Long F, Lei H-T, Bolla JR, Do SV, et al. 2012. Charged amino acids (R83, E567, D617, E625, R669, and K678) of CusA are required for metal ion transport in the Cus efflux system. *J. Mol. Biol.* 422:429--41

118. Su C-C, Long F, Yu EW. 2011. The Cus efflux system removes toxic ions via a methionine shuttle. *Prot. Sci.* 20(6):6--18

119. Su C-C, Long F, Zimmermann MT, Rajashankar KR, Jernigan RL, et al. 2011. Crystal structure of the CusBA heavy-metal efflux complex of *Escherichia coli*. *Nature* 470:558--63

120. Su C-C, Yang F, Long F, Reyon D, Routh MD, et al. 2009. Crystal structure of the membrane fusion protein CusB from *Escherichia coli*. *J. Mol. Biol.* 393(2):342--55

121. Symmons MF, Bokma E, Koronakis E, Hughes C, Koronakis V. 2009. The assembled structure of a complete tripartite bacterial multidrug efflux pump. *Proc. Natl. Acad. Sci. USA* 106(17):7173--78

122. Thieme D, Neubauer P, Nies DH, Grass G. 2008. Sandwich hybridization assay for sensitive detection of dynamic changes in mRNA transcript levels in crude *Escherichia coli* cell extracts in response to copper ions. *Appl. Environ. Microbiol.* 74(24):7463--70

123. Tseng TT, Gratwick KS, Kollman J, Park D, Nies DH, et al. 1999. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol.* 1(1):107--25

124. Van Bambeke F, Balzi E, Tulkens PM. 2000. Antibiotic efflux pumps. *Biochem. Pharmacol.* 60:457--70

125. Veal WL, Nicholas RA, Shafer WM. 2002. Overexpression of the MtrC-MtrD-MtrE efflux pump due to an *mtrR* mutation is required for chromosomally mediated penicillin resistance in *Neisseria gonorrhoeae*. *J. Bacteriol.* 184(20):5619—24

126. Waldron KJ, Rutherford JC, Ford D, Robinson NJ. 2009. Metalloproteins and metal sensing. *Nature* 460:823--30

127. Wells TNC, Scully P, Paravicini G, Proudfoot AEI, Payton MA. 1995. Mechanism of irreversible inactivation of phosphomannose isomerases by silver ions and flamazine. *Biochemistry* 34:7896--903

128. Xue Y, Davis AV, Balakrishnan G, Stasser JP, Staehlin BM, et al. 2008. Cu(I) recognition via cation-π and methionine interactions in CusF. *Nat. Chem. Biol.* 4(2):107--9

129. Yamamoto K, Ishihama A. 2005. Transcriptional response of *Escherichia coli* to external copper. *Mol. Microbiol.* 56(1):215--27

130. Yu EW, Aires JR, McDermott G, Nikaido H. 2005. A periplasmic drug-binding site of the AcrB multidrug efflux pump: a crystallographic and site-directed mutagenesis study. *J. Bacteriol.* 187(19):6804--15

131. Yu EW, McDermott G, Zgurskaya HI, Nikaido H, Koshland DE Jr. 2003. Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump. *Science* 300:976--80

132. Zgurskaya HI, Nikaido H. 2000. Multidrug resistance mechanisms: drug efflux across two membranes. *Mol. Microbiol.* 37(2):219--25

133. Zhang L, Koay M, Maher MJ, Xiao Z, Wedd AG. 2006. Intermolecular transfer of copper ions from the CopC protein of *Pseudomonas syringae*. Crystal structures of fully loaded $Cu^{I}Cu^{II}$ forms. *J. Am. Chem. Soc.* 128:5834—50

CHAPTER III

STRUCTURAL BASIS FOR THE REGULATION OF THE MMPL TRANSPORTERS OF MYCOBACTERIUM TUBERCULOSIS

Delmar, JA, Chou, TH, Wright, CC, Licon, MH, Doh, JK, Radhakrishnan, A, Kumar, N, Lei, HT, Bolla, JR, Rajashankar, KR, Su, CC, Purdy, GE, and Yu, EW. (2015) *J. Biol. Chem.* 290(47): 28559-28574

Introduction

Tuberculosis remains a significant global health problem. Its causative agent, *Mycobacterium tuberculosis* (Mtb), infects more than one third of the world's population (1). The treatment of tuberculosis is notoriously difficult and is recently complicated by the emergence of multiple drug- resistant, extensively drug-resistant, or totally drug-resistant strains (2– 6). The mycobacterial cell wall is unique in composition, and cell wall lipids contribute to Mtb virulence (7–11).

Understanding the molecular mechanisms underlying the biogenesis of the mycobacterial cell wall not only elucidates the basic biology of pathogenic mycobacteria but also identifies potential targets for antimicrobials. Mycolic acids are essential to mycobacterial viability and are incorporated as trehalose dimycolate and arabinogalactan mycolates. Biosynthesis of these respective mycobacterial cell wall lipids is targeted by the first line anti-tuberculosis drugs isoniazid and ethambutol (12-14). Mycobacterial cell wall biosynthesis is facilitated by the MmpL (mycobacterium membrane protein large) transporters (15), which belong to the RND (resistance-nodulation-cell division) superfamily (16). Based on the genomic sequence of *M. tuberculosis* H37Rv (15), this organism harbors 14 different MmpL proteins. MmpL3 is essential, and MmpL4, MmpL5, MmpL7, MmpL8, MmpL10, and MmpL11 are required for full Mtb virulence. Similar to the RND efflux

pumps of Gram-negative bacteria, several of these MmpL transporters appear to work in conjunction with smaller accessory proteins called MmpS (mycobacterial membrane protein small) (7, 17, 18). However, unlike other RND family proteins, the MmpL proteins in *M. tuberculosis* are not believed to export antibiotics (7). Instead, there is strong evidence that these MmpL trans- porters and their MmpS accessory proteins are responsible for shuttling fatty acid and lipid components of the cell wall, such as trehalose monomycolate, sulfolipids, phthiocerol dimycocerosate, diacyltrehalose, monomeromycolyl diacylglycerol, and mycolate wax ester (7–9, 11, 19, 20, 22–25).

Cell wall biogenesis and subsequent remodeling is crucial to the ability of Mtb to establish and maintain infection. It is therefore likely that Mtb carefully controls expression of cell wall lipid biosynthetic enzymes and MmpL lipid transporters. We capitalized on data made available by the Tuberculosis Systems Biology Consortium to begin an in-depth analysis of how *mmpL* and *mmpS* genes are regulated. Currently, ChIP-Seq data for 82 of the 180 Mtb transcription factors are available on the Tuberculosis Database (26 –29). Utilizing these data, we recently showed that the MarR family regulator Rv0678 regulates the *mmpS2-mmpL2*, *mmpS4- mmpL4*, and *mmpS5-mmpL5* genes (30). In addition, we determined the crystal structure of Rv0678, which provided crucial insight into the induction mechanism of Rv0678 (30). Intriguingly, the Rv0678 crystal structure indicated the existence of a fortuitous bound ligand that was identified as 2-palmitoylglycerol ($C_{21}H_{42}O_4$). The induced conformational change leading to substrate-mediated derepression is primarily caused by a rigid body rotational motion of the entire DNA-binding domain of the regulator toward the dimerization domain (30).

In this paper, we report the crystal structures of the TetR family transcriptional

regulators Rv3249c and Rv1816, chosen based upon their predicted regulatory interactions with the *mmpL3* or *mmpL11* loci. Binding of these transcriptional regulators to the promoter and intragenic regions of *mmpL* genes is summarized in Fig. 1. The crystal structures reported here highlight the unique properties of these regulators.

Typically, the TetR family regulators are α-helical dimeric proteins, consisting of a smaller N-terminal DNA-binding domain and a larger C-terminal regulatory domain (31, 32). The N-terminal domains are quite conserved in protein sequences and form a helix-turn-helix motif for DNA binding. However, the C-terminal sequences are poorly conserved, forming ligand-specific binding domains for inducing molecules. Our data indicate that saturated fatty acids are able to bind at the C-terminal ligand-binding domains of Rv3249c and Rv1816. Interestingly, we have recently also found that the *M. tuberculosis* Rv0302 regulator can recognize fatty acid molecules (33).

These findings suggest that fatty acids may be the natural ligands of these regulatory proteins. We used fluorescence polarization and EMSA to demonstrate that these proteins regulate multiple *mmpL* genes. These results emphasize the complexity of the *mmpL* regulatory system and a novel mechanism by which the bacterium can sense metabolic state to modulate cell wall lipid biosynthesis and transport to maintain homeostasis and promote virulence.

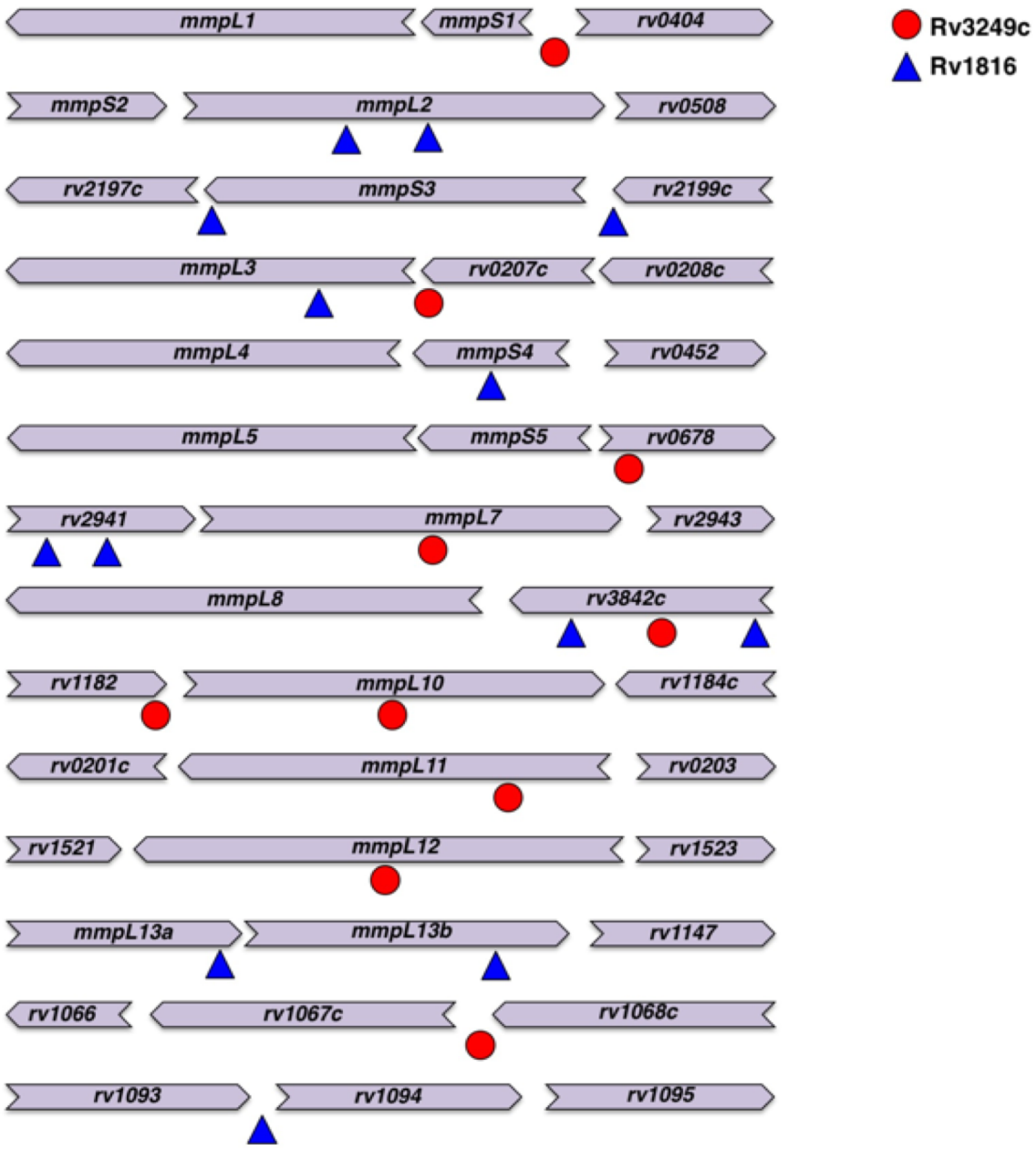

Figure 1. *Schematic depiction of transcription factor binding sites in the mmpL genes of interest. ChIP-Seq data for each transcription factor was obtained from Tuberculosis Database. In these experiments, FLAG-tagged transcription factors were episomally expressed in Mtb under the control of an anhydrotetracycline-inducible promoter (29). The colored shapes corresponding to each transcription factor (red circle, Rv3249c; blue triangle, Rv1816) are placed at the putative binding sites.*

Experimental Procedures

Cloning of rv3249c and rv1816

The *rv3249c* ORF from genomic DNA of *M. tuberculosis* strain H37Rv was amplified by PCR using the primers 5'-CTTTAAGAAGGAGATATACCATGGTGAGCACTCCAAGCGCTAC-3' and 5'-GATCCTCAGTGATGATGGTGGTGATGGGGGACATTGATCACACCGTG-3' to generate a product that encodes a Rv3249c recombinant protein with a His_6 tag at the C terminus. The corresponding PCR product was digested with NcoI and BamHI, extracted from the agarose gel, and inserted into pET15b as described by the manufacturer (Merck). The recombinant plasmid (pET15bΩ*rv3249c*) was transformed into DH10b cells, and the transformants were selected on LB agar plates containing 100 μg/mL ampicillin. The presence of the correct *rv3249c* sequence in the plasmid construct was verified by DNA sequencing.

The procedures for producing the recombinant plasmid pET15bΩ*rv1816* were the same as those for pET15bΩ*rv3249c*. The *rv1816* ORF from genomic DNA of *M. tuberculosis* strain H37Rv was amplified by PCR using the primers 5'-CTTTAAGAAGGAGATATACCATGTTGTGTCAGACTTGCCGCGTG-3' and '5-GATCCTCAGTGATGATGGTGGTGATGCTCGGCCAGCACGGCCAC-3' to generate a product that encodes the Rv1816 recombinant protein with a His_6 tag at the C terminus.

Expression and purification of Rv3249c and Rv1816

Briefly, the full-length Rv3249c protein containing a His_6 tag at the C terminus was overproduced in *E. coli* BL21(DE3) cells possessing pET15bΩ*rv3249c*. The cells were

grown in 6 L of LB medium with 100 μg/mL ampicillin at 37 °C. When the A_{600} reached 0.5, the culture was treated with 0.2 mM isopropyl-β-D-thiogalactopyranoside to induce Rv3249c expression, and cells were harvested within 3 h. The collected bacterial cells were suspended in 100 mL of ice-cold buffer containing 20 mM Na-HEPES (pH 7.2), 200 mM NaCl, 10 mM $MgCl_2$, and 0.2 mg of DNase I (Sigma-Aldrich). The cells were then lysed with a French pressure cell. Cell debris was removed by centrifugation for 45 min at 4 °C and 20,000 rev/min. The crude lysate was filtered through a 0.2 μm membrane and was loaded onto a 5-mL Hi-Trap Ni^2-chelating column (GE Healthcare Biosciences) pre-equilibrated with 20 mM Na-HEPES (pH 7.2) and 200 mM NaCl. To remove unbound proteins and impurities, the column was first washed with six column volumes of buffer containing 50 mM imidazole, 250 mM NaCl, and 20 mM Na-HEPES (pH 7.2). The Rv3249c protein was then eluted with four column volumes of buffer containing 300 mM imidazole, 250 mM NaCl, and 20 mM Na-HEPES (pH 7.2). The purity of the Rv3249c protein (>95%) was judged using 12.5% SDS-PAGE stained with Coomassie Brilliant Blue. The purified protein was then concentrated to 10 mg/mL.

For the His_6 SeMet-Rv3249c protein expression, a 10 mL LB broth overnight culture containing *E.coli* BL21(DE3)/pET15bΩ*rv3249c* cells was transferred into 60 mL of LB broth containing 100 μg/mL ampicillin and grown at 37 °C. When the A_{600} value reached 1.2, cells were harvested by centrifugation at 5,000 *g* for 10 min and then washed two times with 10 mL of M9 minimal salt solution. The cells were resuspended in 60 mL of M9 medium and then transferred into a 6 L pre-warmed M9 solution containing 100 μg/mL ampicillin. The cell culture was incubated at 37 °C with shaking. When the A_{600} reached 0.4, 100 mg/L of lysine, phenylalanine, and threonine, 50 mg/L isoleucine, leucine, and valine, and 60 mg/L of

L-selenomethionine were added. The culture was induced with 0.2 mM isopropyl-β-D-thiogalactopyranoside after 15 min. The cells were then harvested within 3 h after induction. The procedures for purifying SeMet-Rv3249c were identical to those of the native protein.

The full-length Rv1816 protein containing a His_6 tag at the C terminus was overproduced in *E. coli* BL21(DE3) cells possessing pET15bΩ*rv1816*. The procedures for expressing and purifying the Rv1816 protein were the same as those for Rv3249c. The purity of the Rv1816 protein (>95%) was judged using 12.5% SDS-PAGE stained with Coomassie Brilliant Blue. The purified protein was then concentrated to 10 mg/mL.

Crystallization of Rv3249c and Rv1816

The His_6 Rv3249c crystals were grown in microcentrifuge tubes (Fisher Scientific). Briefly, 10 μL aliquots of protein solution containing 20 mg/mL Rv3249c protein in 20 mM Na-HEPES (pH 7.5), 250 mM NaCl, and 300 mM imidazole were incubated at 4 °C. Crystals of Rv3249c grew to a full size within 2 weeks. Typically, the dimensions of the crystals were 0.05 mm x 0.05 mm x 0.1 mm. Cryoprotection was achieved by raising the glycerol concentration stepwise to 40% with a 5% increment in each step. Crystals of SeMet-Rv3249c were obtained using the same procedures as for the native Rv3249c crystals. Typically, the dimensions of the crystals were 0.05 mm x 0.05 mm x 0.1 mm.

The His_6 Rv1816 crystals were grown at room temperature in 24-well plates (Hampton Research, Aliso Viejo, CA) using hanging drop vapor diffusion with the following procedures. A 1 μL protein solution containing 10 mg/mL Rv1816 protein in 20 mM Na-HEPES (pH 7.5), 250 mM NaCl, and 300 mM imidazole was mixed with 1 μL of reservoir solution containing 0.05 M sodium cacodylate (pH 6.5), 2 M ammonium sulfate, and 10 mM

magnesium sulfate. The resultant mixture was equilibrated against 500 μL of the reservoir solution. Hexagonal crystals appeared within 2 days and grew to full size in the drop within 1 week. Typically, the dimensions of the crystals were 0.05 mm x 0.05 mm x 0.1 mm. Cryoprotection was achieved by raising the ammonium sulfate concentration stepwise to 3.5 M.

Data collection, structural determination, and refinement

All diffraction data were collected at 100 K at Beamline 24ID-C located at the Advanced Photon Source, using a Pilatus 6M detector (Dectris Ltd., Switzerland). Diffraction data were processed using DENZO and scaled using SCALEPACK (34). The Rv3249c crystals belong to the space group C2 (Table 1). Based on the molecular weight of Rv3249c (22.8 kDa), six Rv3249c molecules are expected in the asymmetric unit with a solvent content of 62.96%. Within the asymmetric unit of SeMet-Rv3249c, 12 selenium sites (two per monomer) were identified using SHELXC and SHELXD (35) as implemented in the HKL2MAP package (36). The full-length Rv3249c protein contains two methionine residues; both were identified as SeMet sites in each protomer. Single anomalous dispersion was employed to obtain experimental phases using the program MLPHARE (37, 38). The resulting phases were then subjected to density modification and NCS averaging using the program PARROT (39) using the native structure factor amplitudes. The SeMet sites were also used to trace the molecules by anomalous difference Fourier maps, where we could ascertain the proper registry of SeMet residues. The initial model, which consists of three dimeric molecules, was constructed manually using program Coot (40). Then the model was refined using PHENIX (41), leaving 5% of reflections in the Free-R set.

Table 1. Data collection, phasing and structural refinement statistics of Rv3249c and Rv1816.

Data set	Rv3249c Native	SeMet-Rv3249c	Rv1816
Data collection			
Wavelength (Å)	0.9792	0.9792	0.9792
Space group	C2	C2	$P2_12_12_1$
Cell constants (Å)			
a	233.5	232.0	40.5
b	75.3	75.4	86.5
c	125.0	121.0	129.9
α, β, γ (°)	90, 121.6, 90	90, 121.1, 90	90, 90, 90
Resolution (Å)	3.59 (3.72-3.59)	3.40 (3.52-3.40)	2.00 (2.07-2.00)
Completeness (%)	95.7 (97.6)	98.4 (96.1)	99.9 (100.0)
Total reflections	491,698	2,356,230	1,011,787
Unique reflections	22,121	24,714	31,661
Redundancy	3.5 (3.6)	5.8 (5.1)	6.3 (6.5)
R_{merge} (%)	11.4 (39.1)	5.0 (41.6)	6.4 (41.3)
$\langle I/\sigma(I)\rangle$	11.0 (4.2)	27.9 (3.5)	24.1 (5.4)
Phasing			
Number of sites		12	
Resolution used (Å)		4.0	
Phasing power (acentric/centric)		1.0/0.69	
R_{Cullis} (acentric/centric)		0.88/0.90	
Figure of merit (acentric/centric)		0.43/0.20	
Refinement			
Resolution (Å)	40-3.59		40-2.00
R_{work} (%)	17.70		17.75
R_{free} (%)	27.82		22.16
Overall B-factor ($Å^2$)	51.37		55.68
Rms deviations			
Bond (Å)	1.676		1.392
Angles (^{o})	0.013		0.013
Ramachandran analysis			
Most favored (%)	97		94.7
Allowed (%)	2.5		5.2
Generously allowed (%)	0.5		0.1
Disallowed (%)	0		0.0

Iterations of refinement using PHENIX (41) and CNS (42) and model building in Coot (40) lead to the current model, which consists of three dimeric structures with a total number of 1,160 residues (Table 1).

The Rv1816 crystal belongs to the space group $P2_12_12_1$, resulting in a single Rv1816 dimer (25.4 kDa) in the asymmetric unit with a solvent content of 51.78%. The structure of Rv1816 was determined using molecular replacement. The partial structure of the TetR family protein SCO4313 (residues 61–229) (Protein Data Bank code 2OI8) was utilized as a template. The resulting phases were then subjected to density modification using the program RESOLVE (43). The remaining part of the model was manually constructed using the program Coot (40). Model was then refined using PHENIX (41), leaving 5% of reflections in the Free-R set. Iterations of refinement using PHENIX (41) and CNS (42) and model building in Coot (40) lead to the current model of the Rv1816 dimer, consisting of a total number of 449 residues (Table 1).

Identification of fortuitous ligands

To identify the nature of the bound ligand in crystals of Rv3249c, we used GC-MS. The Rv3249c crystals were extensively washed with the crystallization buffer and transferred into deionized water. The mixture was then incubated at 100 °C for 5 min, and then chloroform was added into the mixture to a final concentration of 80% (v/v) to denature the protein and allow for the extraction of ligand. GC-MS analysis indicated that the mass of the bound ligand was *n*-hexadecanoic acid, also called palmitic acid.

We used the same method to identify the bound ligand in crystals of Rv1816. GC-MS analysis indicated that the mass of the bound ligand corresponded to isopropyl dodecanoate, also called isopropyl laurate.

Isothermal titration calorimetry for ligand binding

We used isothermal titration calorimetry (ITC) to determine the binding affinity of palmitic acid to the purified Rv3249c regulator. Measurements were performed on a VP-Microcalorimeter (MicroCal, Northampton, MA) at 25 °C. Before titration, the protein was thoroughly dialyzed against buffer containing 10 mM sodium phosphate (pH 7.2), 100 mM NaCl, and 0.001% DDM. The protein concentration was determined using the Bradford assay. The dimeric Rv3249c sample was then adjusted to a final concentration of 10 μM. The ligand solution containing 500 μM palmitic acid, 10 mM sodium phosphate (pH 7.2), 100 mM NaCl, and 0.001% DDM was used as the titrant. Binding experiments were carried out with the protein solution (1.4 mL) in the cell and the ligand solution as the injectant. Thirty injections of 10 μL each of the ligand solution were used for data collection.

Injections occurred at intervals of 240 s, and the duration time of each injection was 20 s. Heat transfer (μcal/s) was measured as a function of elapsed time (s). The mean enthalpies measured from injection of the ligand in the buffer were subtracted from raw titration data before data analysis with ORIGIN software (MicroCal). Titration curves were fitted by a non-linear least squares method to a function for the binding of a ligand to a macromolecule. Nonlinear regression fitting to the binding isotherm provided us with the equilibrium binding constant ($K_A = 1/K_D$) and enthalpy of binding (ΔH). Based on the values of K_A, the change in free energy (ΔG) and entropy

(ΔS) were calculated with the equation: $\Delta G = -RT \ln K_A = \Delta H - T\Delta S$, where T is 273 K and R is 1.9872 cal/K per mol. Calorimetry trials were also carried out in the absence of Rv3249c in the same experimental conditions. No change in heat was observed in the injections throughout the experiment.

ITC was also used to determine the binding affinities of palmitate and laurate to the purified Rv1816 regulator. The Rv1816 protein was thoroughly dialyzed against buffer containing 10 mM sodium phosphate (pH 7.2), 100 mM NaCl, and 0.001% DDM. The dimeric Rv1816 sample was then adjusted to a final concentration of 10 μM. The ligand solution contained 400 μM palmitic or lauric acid, 10 mM sodium phosphate (pH 7.2), 100 mM NaCl, and 0.001% DDM. The procedures for these binding experiments were identical to those for the Rv3249c protein.

Fluorescence polarization assay for DNA binding

Fluorescence polarization assays were used to determine the affinities for DNA binding by Rv3249c and Rv1816, respectively. All oligodeoxynucleotides and their corresponding fluorescein-labeled oligodeoxynucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). The sequences of these oligodeoxynucleotides are indicated in Table 2. In short, the fluoresceinated dsDNA was prepared by annealing the oligodeoxynucleotide and its corresponding fluorescein-labeled oligodeoxynucleotide together. Fluorescence polarization experiment was done using a DNA binding solution containing 10 mM sodium phosphate (pH 7.2), 100 mM NaCl, 2.5 nM fluoresceinated DNA, and 1 μg of poly(dI-dC) as nonspecific DNA. The protein solution containing 500 nM dimeric Rv3249c or Rv1816, and 2.5 nM of the respective fluoresceinated DNA was titrated

into the DNA binding solution until the millipolarization became unchanged. All measurements were performed at 25 °C using a PerkinElmer LS55 spectrofluorometer equipped with a Hamamatsu R928 photomultiplier. The excitation wavelength was 490 nm, and the fluorescence polarization signal (in ΔP) was measured at 525 nm. Each titration point recorded was an average of 15 measurements. The data were analyzed using the equation, $p = \{(P_{bound} - P_{free})[protein]/(K_D + [protein])\} + P_{free}$, where P is the polarization measured at a given total protein concentration, P_{free} is the initial polarization of free fluorescein-labeled DNA, P_{bound} is the maximum polarization of specifically bound DNA, and [protein] is the protein concentration. The titration experiments were repeated for three times to obtained the average K_D value. Curve fitting was accomplished using the program ORIGIN (OriginLab Corporation, Northampton, MA).

Table 2. Affinity for DNA binding by Rv3249c or Rv1816.

Transcription Factor	DNA sequence	Location	K_d (nM)
Rv3249c	5′-TTTTGACCTCGCCGTAAACGATGTAAACG-3′ 5′-F-CGTTTACATCGTTTACGGCGAGGTCAAAA-3′[a]	*mmpS1-mmpL1*	18.7 ± 9.5
	5′-CGTGACATCGACCGTTTCGATGTTCGG-3′ 5′-F-CCGAACATCGAAACGGTCGATGTCACG-3′[a]	*mmpL11*	5.6 ± 1.0
	5′-AATGAATGTCAACACATTTGACATTTCAG-3′ 5′-F-CTGAAATGTCAAATGTGTTGACATTCATT-3′[a]	*rv1067c*	26.9 ± 4.9
Rv1816	5′-CTTCGCGCGGGTCGCGCGGCT-3′ 5′-F-AGCCGCGCGACCCGCGCGAAG-3′[a]	*mmpL13b*	4.3 ± 0.4
	5′-ACAGATTTCGTGAAATCGGG-3′ 5′-F-CCCGATTTCACGAAATCTGT-3′[a]	*rv1094*	6.9 ± 0.7

[a] F denotes the fluorescein that was covalently attached to the 5′ end of the oligonucleotide (reversed) by a hexamethylene linker.

Electrophoretic mobility shift assay

Probes were amplified from the H37Rv genome using the primers listed in Table 3. All probes were labeled with digoxigenin using the Roche DIG gel shift kit. For EMSA analysis, 12 nM Dig-labeled probe and the indicated micromolar concentrations of protein were incubated for 45 min at room temperature in the Roche binding buffer modified by the addition of 0.25 mg/mL herring sperm DNA, and 0.75 mg/mL poly(d[I-C]). For ligand competition assays, stock solutions of fatty acids were made in DMSO and a solvent control reaction included at the highest concentration of DMSO. All reactions were resolved on a 6% native polyacrylamide gel in TBE buffer and transferred to nylon membrane and Dig-labeled DNA-protein complexes detected following the manufacturer's recommendations. Chemiluminescent signals were acquired using an ImageQuant LAS 4000 (GE).

Table 3. Primers for EMSA probes.

Rv1816–1F	5′-CTCGGTTGTGTCAGACTTGC-3′
Rv1816–1R	5′-CAACAGCTCATCACGACTGG-3′
Rv1816–2F	5′-TGGGCCTTGCTATACGGTAG-3′
Rv1816–2R	5′-GCAGAAAGCACTTTGTGACG-3′
mmpL3.1816F	5′-GGCAACACGGTGATAGACAG-3′
mmpL3.1816R	5′-GAAGGCTACGACACCGAGAC-3′
Rv0204.1816F	5′-GGATGACAAGTCGCAGGTTT-3′
Rv0204.1816R	5′-CATCTCGGCGATGTTGATCT-3′
mmpl11.1816F	5′-ATCTCGTGTGACGAGTCGAA-3′
mmpl11.1816R	5′-GTCAGCGTTAGTGCACTGGT-3′
mmpL7.1816F	5′-AGTACGTCGTCGCCTTTCTC-3′
mmpL7.1816R	5′-ATAATTGATGGCGGGGATTC-3′
kasA.1816F	5′-TCCAAGTGAGTCAGCCTTCC-3′
kasA.1816R	5′-ACCGGATCCTTGAGGTGAC-3′
mmpS3.1816F	5′-GTCCATCGTTGGAGTTCGAT-3′
mmpS3.1816R	5′-CGGTGACTTACGTCGATGC-3′
rv1094.1816F	5′-TCGAGGAGTGGAGTCTGGTC-3′
rv1094.1816R	5′-GACGTAGTCGTGGGCGAAC-3′
mmpL3.3249cF	5′-ATGACCCCGATTACGATGAA-3′
mmpL3.3249cR	5′-GAGGACATCGCTGGTGTTTT-3′
mmpL11.3249cF	5′-CGGTAGCGCATCAGGATGAA-3′
mmpL11.3249cR	5′-CGCTGGAACCTGCCTATCAT-3′
rv3249c.3249cF	5′-AGAGCACATTGGCGTCGTTA-3′
rv3249c.3249cR	5′-CGGCAACTTTTACGAAGCGT-3′
mmpS1.3249cF	5′-TGATGTGAGCCAAACACCGA-3′
mmpS1.3249cR	5′-TTTGACCTCGCCGTAAACGA-3′
rv1067.3249cF	5′-GAATTGCTCGTGCGATCCGG-3′
rv1067.3249cR	5′-GATCCGGTGTGACGACGTATC-3′

Results and Discussion

Overall structure of Rv3249c

The crystal structure of the 211-amino acid Rv3249c protein was determined to a resolution of 3.59 Å using single anomalous dispersion (Table 1 and Fig. 2). Six molecules of Rv3249c are found in the asymmetric unit, which assemble as three independent dimers. Superimposition of these six Rv3249c molecules gives root mean square deviations between 0.8 and 1.0 Å over 190 Cα atoms, indicating that their conformations are nearly identical to each other.

Like other members of the TetR family regulators (31, 32), the structure of Rv3249c suggests that this regulator is an all α-helical protein. Each subunit of Rv3249c dimer is composed of 10 α-helices (α1-α10 and α1'-α10', respectively) (Fig. 3). These helices are designated as α1 (16–36), α2 (49–53), α3 (57–64), α4 (67–92), α5 (97–114), α6 (116–123), α7 (128– 134), α8 (138 –155), α9 (162–181), and α10 (190 –205). The overall architecture can be divided into two domains. Residues 16 –77, comprising α1, α2, and α3, and the N-terminal end of α4, form the DNA-binding domain. Helices α2 and α3 exhibit the classical helix-turn-helix motif, typical of these regulators. Residues 68 –205, comprising the C-terminal end of α4 through α10, form the ligand-binding domain of Rv3249c. Helices α4, α5, α8, and α9 form an antiparallel four-helix bundle, which creates a large ligand-binding cavity to accommodate for the inducing ligands. Helices α6 and α7, and the loop connecting these two short helices create the floor of this ligand-binding site. It should be noted that a typical TetR regulator normally utilizes a single helix, which runs horizontally to form the bottom of the ligand-binding site. Rv3249c is distinct in that it employs two inclined helices to create a V-shaped bottom, allowing it to enlarge the volume of the binding site.

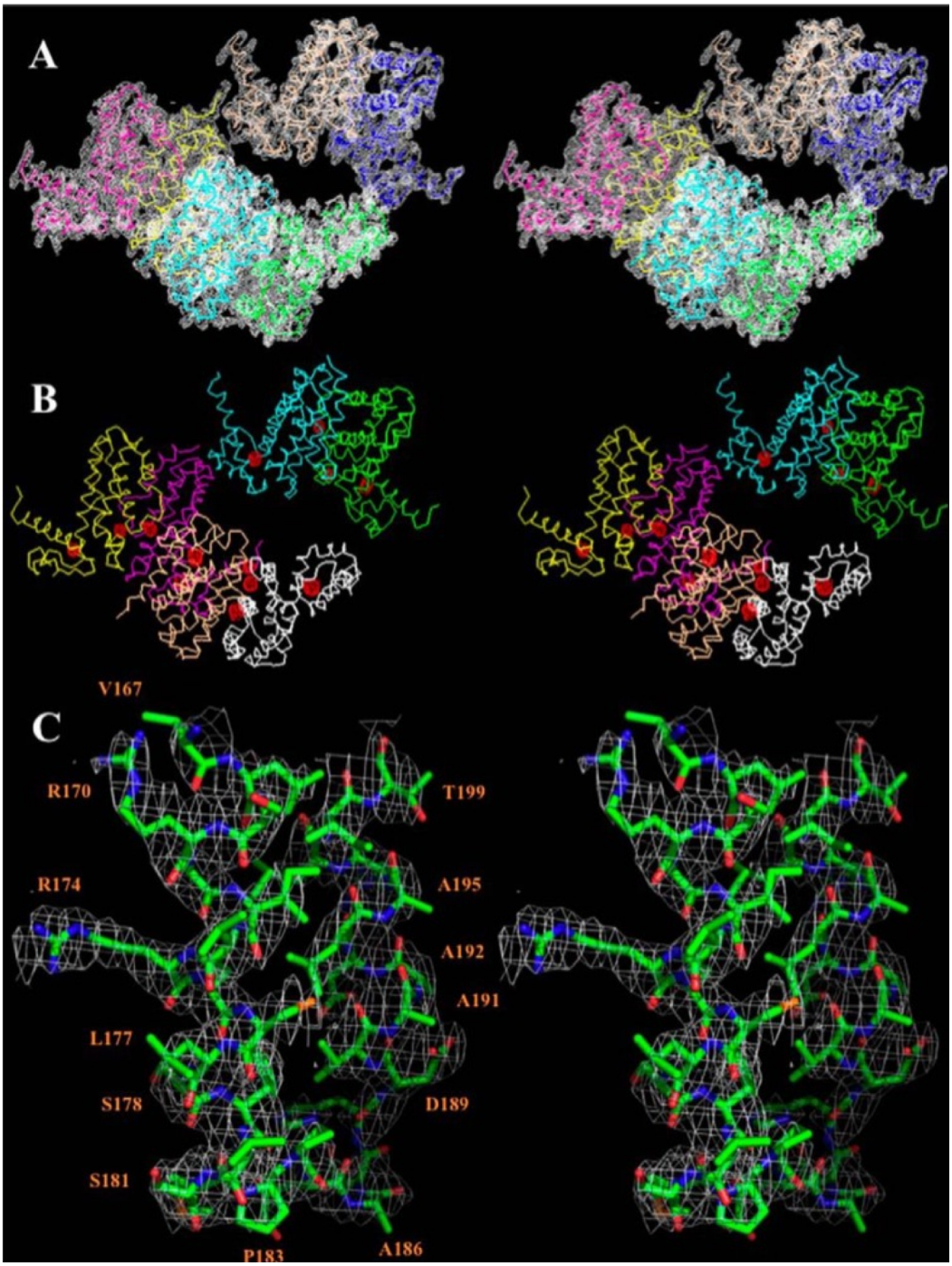

Figure 2. *Stereo view of the experimental electron density maps of Rv3249c at a resolution of 3. 59 Å. A, the electron density maps are contoured at 1.0 σ. The Cα traces of the three Rv3249c dimers in the asymmetric unit are in slate, wheat, green, cyan, yellow, and magenta. B, anomalous signals of the 12 selenium sites (contoured at 5 σ) found in the asymmetric unit are colored red. The Cα traces of the three Rv3249c dimers in the asymmetric unit are in green, cyan, white, light pink, magenta, and yellow. C, representative section of electron density in the vicinity of helices α9 and α10. The solvent-flattened electron density (40–3.59 Å) is contoured at 1.0 σ and superimposed with the final refined model (green, carbon; red, oxygen; blue nitrogen; yellow, sulfur).*

Indeed, each subunit of Rv3249c possesses a very large cavity, which was estimated to be 450 $Å^3$ within the C-terminal ligand binding domain. Based on this observation, we expect that this regulator can accommodate a large inducing ligand.

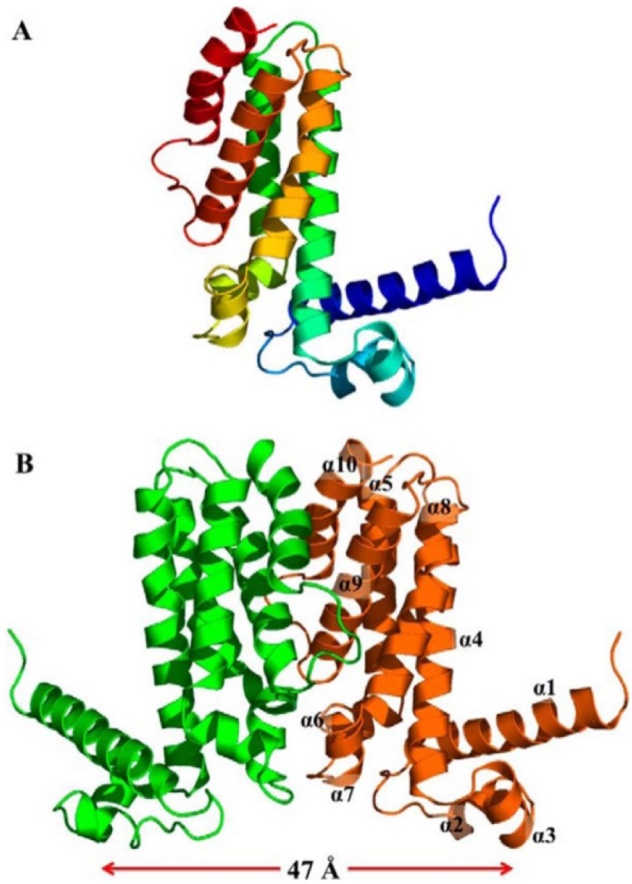

Figure 3. *Structure of the M. tuberculosis Rv3249c regulator. A, ribbon diagram of a protomer of Rv3249c. The molecule is colored using a rainbow gradient from the N terminus (blue) to the C terminus (red). B, ribbon diagram of the Rv3249c dimer. Each subunit of Rv3249c is labeled with a different color (orange and green). The figure was prepared using PyMOL.*

The smaller N-terminal domain of Rv3249c shares considerable structural similarities with other TetR family members. However, this domain also presents some noticeable differences. Helix α1, consisting of 21 amino acids, is relatively long compared with other members of the family. This helix tilts upward by 15° in relation to the horizontal surface, which is perpendicular to the vertical plane formed by the dimer interface. This orientation facilitates the two N-terminal domains within the dimer to shift away from each other, leaving a relatively large gap between the two DNA recognition helices α3

and α3'. Helices α3 and α3' play a key role in the binding of cognate DNA. Because the distance between two consecutive major grooves of a B-form DNA is ~34 Å, the center to center distance between helices α3 and α3' of the regulator has to be about 34 Å to bind the DNA. In Rv3249c, this center to center distance, according to the separation between Cα atoms of Y61 and Y61', was measured to be 47 Å, suggesting that, in this conformation, the regulator cannot bind B-DNA, and this conformation should correspond to the induced form.

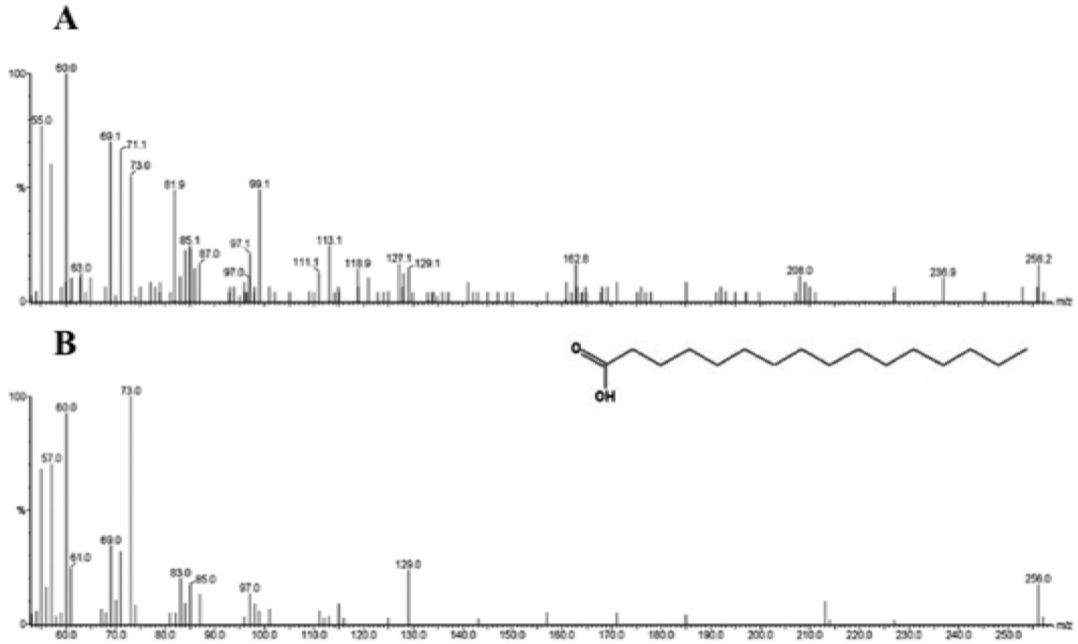

Figure 4. *Identification of the fortuitous ligand of Rv3249c by GC-MS. A, electron ionization spectrum of the strongest GC peak at 10.45 min. B, GC-MS spectrum of n-hexadecanoic acid from the internal GC-MS library. The ligand was identified as palmitic acid.*

The relatively large center to center distance also indicated that the crystallized Rv3249c protein has an attached ligand. As expected, a large extra electron density was found within the C-terminal regulatory domain of the ligand-binding site of each subunit of Rv3249c, indicating the existence of a fortuitous bound ligand co-purified and co-crystallized with this regulator. To identify the unknown bound ligand, GC-MS was used to study these

Rv3249c crystals (Fig. 4). The data suggest that the fortuitous ligand is palmitic acid, also called hexadecanoic acid, a saturated fatty acid containing 16 carbons with the molecular formula $C_{16}H_{32}O_2$. Because this fatty acid was found to be co-purified and co-crystallized with the Rv3249c regulator, fatty acids may be the natural substrates of this protein.
The bound palmitic acid is completely buried in the Rv3249c binding cavity (Fig. 5). Extensive hydrophobic interactions are found to anchor this acid. Within 5 Å of the bound ligand, there are at least 22 amino acids, including Ala-75, Leu-78, Ala-79, Leu-82, Leu-83, Val-86, Phe-101, Gly-104, Phe-105, Phe-108, Leu-117, Val-118, Leu-130, Leu-133, Ile-141, Ala-145, Leu-149, Ile-172, Val-173, Cys-176, Leu-177, and Val-180, involved to secure the binding (Fig. 5). Many of these residues are hydrophobic in nature.

Overall structure of Rv1816

The crystal structure of the 234-amino acid protein Rv1816 was determined to a resolution of 2.00 Å (Table 1 and Fig. 6). Because *M. tuberculosis* Rv1816 shares 31% protein sequence identity with *Streptomyces coelicolor* SCO4313 (Protein Data Bank code 2OI8), we utilized the existing structure of SCO4313 from the Protein Data Bank as a search model for molecular replacement to obtain the crystal structure of Rv1816. Two molecules of Rv1816, which assemble as a dimer, are found in the asymmetric unit. Similar to Rv3249c, the structure of Rv1816 can be divided into N-terminal DNA-binding and C-terminal ligand-binding domains. This regulator is also an all α-helical protein, composed of 10 α-helices (α1–α10 and α1'–α10', respectively) (Fig. 7): α1 (13– 31), α2 (39 – 46), α3 (50 –56), α4 (60 – 85), α5 (91–108), α6 (110 –117), α7 (128 –152), α8 (164 –177), α9 (183–205), and α10 (216 –232).

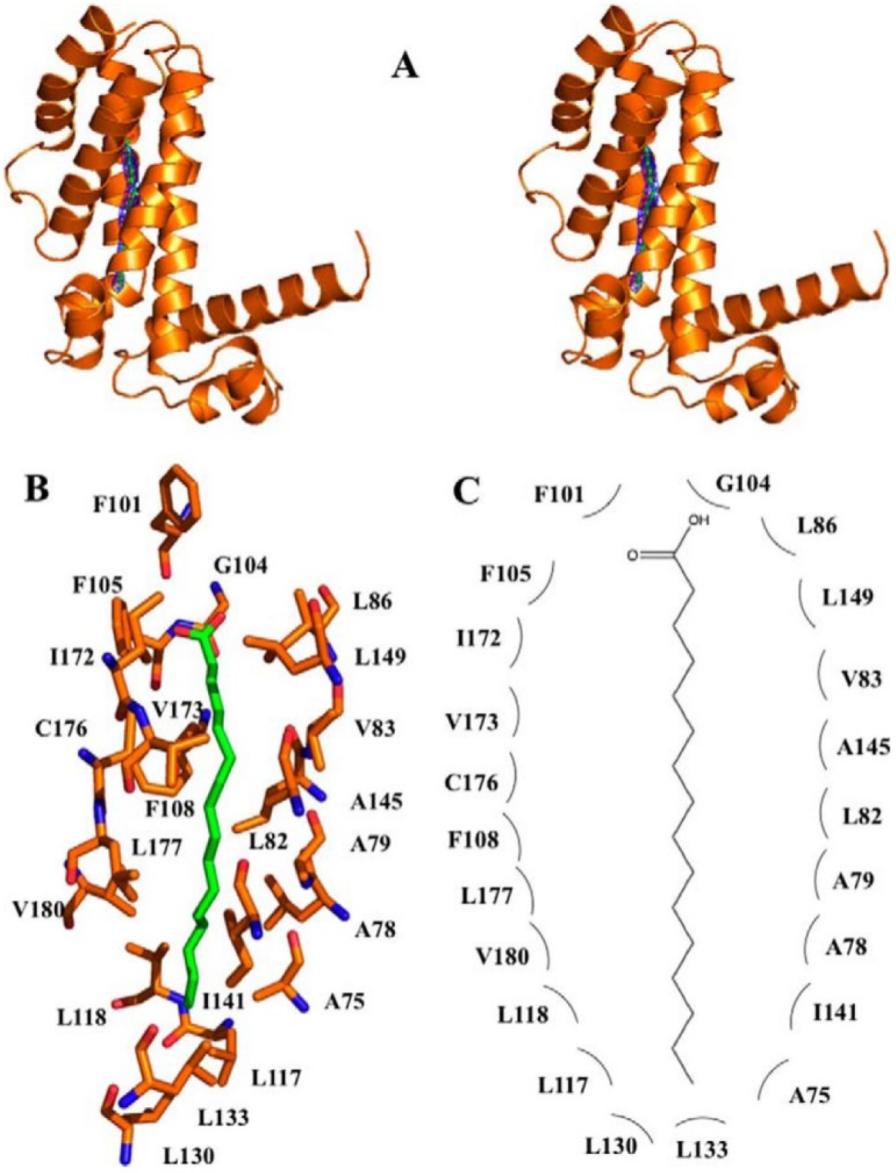

Figure 5. *Simulated annealing electron density maps and the palmitate binding site. A, stereo view of the simulated annealing electron density map of the bound palmitate within the Rv3249c protomer (the orientation corresponds to the side view of Fig. 1B). The bound palmitate is shown as a stick model (green, carbon; red, oxygen). The simulated annealing $2F_o - F_c$ electron density map is contoured at 1.0 σ (blue mesh). The right subunit of Rv3249c is shown as orange ribbons. B, the palmitate binding site. Amino acid residues within 5.0 Å from the bound palmitate (green, carbon; red, oxygen) are included. The side chains of selected residues from the right subunit of Rv3249c in Fig. 1B are shown as orange sticks (orange, carbon; blue, nitrogen; red, oxygen). C, schematic representation of the Rv3249c and palmitate interactions. Amino acid residues within 5.0 Å from the bound palmitate are included.*

The structure of Rv1816 is unique among the TetR family regulators. In the C-terminal ligand-binding domain, helices α9 and α10 of one subunit of Rv1816, which run anti-parallel to each other, interact with helices α9’ and α10’ of the next subunit, forming a four-helix bundle at the subunit interface to secure the dimerization state. A tunnel-like cavity is formed at the cleft between helices α7, α8, α9, and α10 and the random loop connecting α9’ and α10’ of the second subunit. Unlike Rv3249c, a single helix, α6, forms the floor of this cavity, which potentially creates a binding tunnel for inducing ligands. Helix α8, which orients horizontally and nearly perpendicular to the 2-fold vertical axis of the dimer, generates a unique architecture of the Rv1816 regulator. This feature has not been found in the existing structures of the TetR family regulators including TetR (44, 45), QacR (46, 47), CprB (48), EthR (49, 50), CmeR (51, 52), AcrR (53), SmeT (54), Rv3066 (55), Rv1219c (56), and Rv3249c. Based on its location, helix α8, together with the loop region between helices α9’ and α10’, seems to lengthen and enlarge the binding cavity

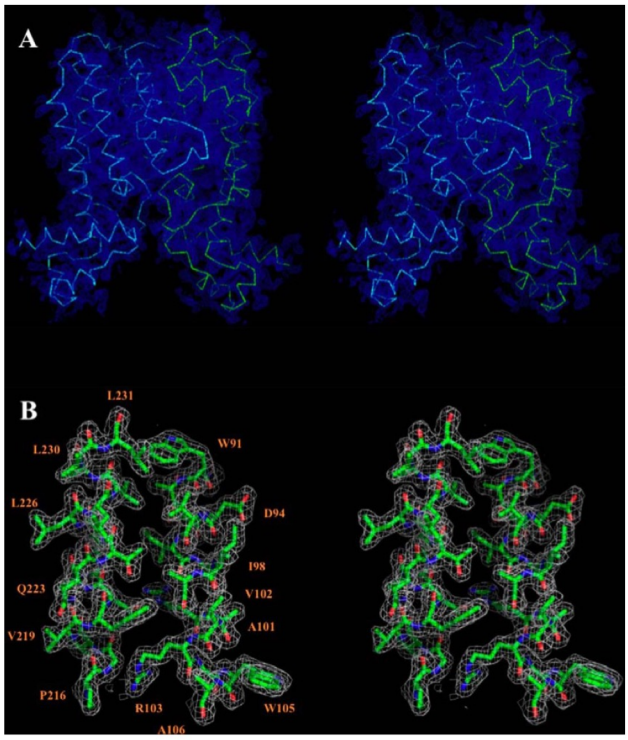

Figure 6. *Stereo view of the electron density maps of Rv1816 at a resolution of 2. 00 Å. A, the electron density maps are contoured at 1.0 σ. The Cα traces of the Rv1816 dimer in the asymmetric unit are in green and cyan. B, representative section of electron density in the vicinity of helices α5 and α10. The solvent-flattened electron density (40–2.00 Å) is contoured at 1.0 σ and superimposed with the final refined model (green, carbon; red, oxygen; blue, nitrogen; yellow, sulfur).*

to accommodate for larger ligands. Indeed, it was found that the loop residues as well as residues located within helix α8 are involved in ligand binding. The volumes of the ligand binding cavities of the C-terminal regulatory domains of the two subunits within the Rv1816 dimer are 500 $Å^3$ and 450 $Å^3$ (corresponding to the right and left subunits of Rv1816 in Fig. 7*B*, respectively).

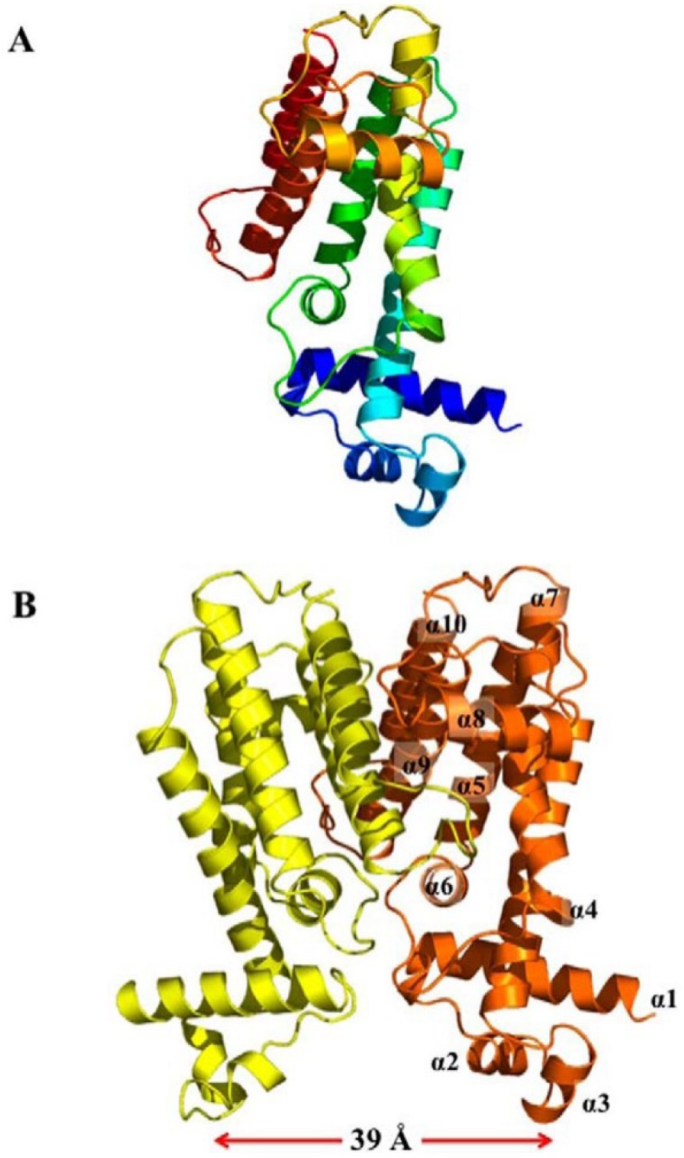

Figure 7. *Structure of the M. tuberculosis Rv1816 regulator. A, ribbon diagram of a protomer of Rv1816. The molecule is colored using a rainbow gradient from the N terminus (blue) to the C terminus (red). B, ribbon diagram of the Rv1816 dimer. Each subunit of Rv1816 is labeled with a different color (orange and yellow).*

Within the N-terminal DNA-binding domain, helix α1 of Rv1816 is relatively long at 19 amino acids. However, in direct contrast to the structure of Rv3249c, this helix tilts downward by 10° with respect to the horizontal plane. The center-to-center distance between helices α3 and α3' of the regulator is 39 Å, defined by the separation between residues Tyr-54 and Tyr-54'. Such distances in the apo form of TetR (44), QacR (46), and AcrR (53) are 35, 39, and 42 Å,

respectively. This spacing indicates that the crystal structure of Rv1816 may represent its ligand-free conformational state. Surprisingly, a large extra electron density appears in the elongated binding pocket of one subunit of the Rv1816 protein, suggesting that Rv1816 is bound by a ligand. However, the other binding pocket in the next sub-unit of the regulator is still empty. As in the case of Rv3249c, GC-MS was employed to identify this fortuitous ligand. The data indicate that Rv1816 was co-purified and co-crystallized with the saturated fatty acid ester isopropyl laurate, also called propan-2-yl dodecanoate, with the molecular formula $C_{15}H_{30}O_2$ (Fig. 8). Again, these data support the role of fatty acids as the natural

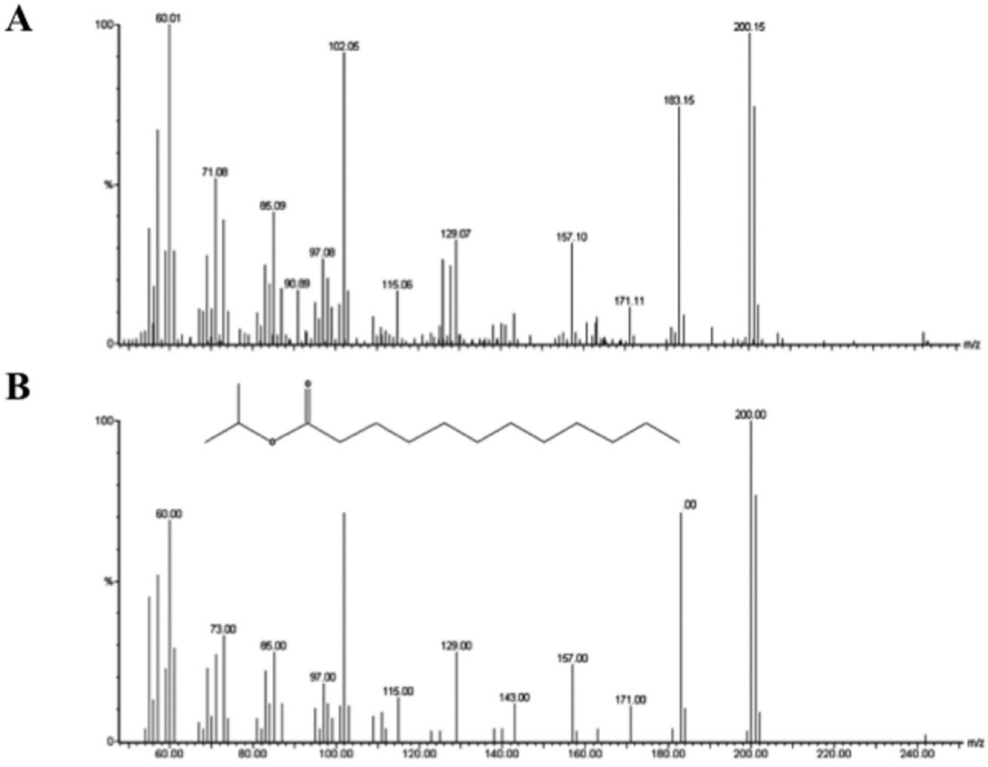

Figure 8. *Identification of the fortuitous ligand of Rv1816 by GC-MS. A, electron ionization spectrum of the strongest GC peak at 8.58 min. B, GC-MS spectrum of propan-2-yl dodecanoate from the internal GC-MS library. The ligand was identified as isopropyl laurate.*

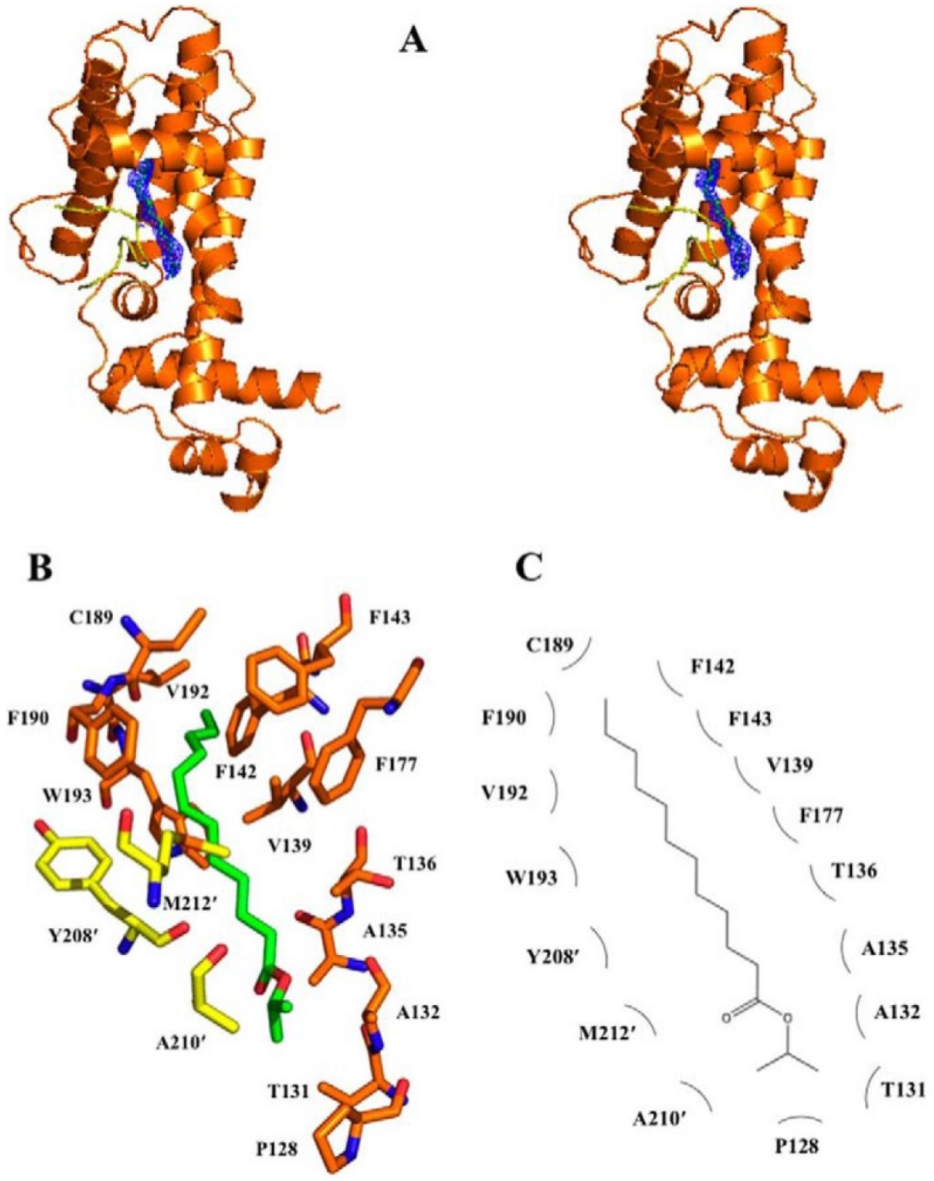

Figure 9. *Simulated annealing electron density maps and the isopropyl laurate binding site. A, stereo view of the simulated annealing electron density map of the bound isopropyl laurate within the Rv1816 protomer (the orientation corresponds to the side view of Fig. 7B). The bound isopropyl laurate is shown as a stick model (green, carbon; red, oxygen). The simulated annealing $2F_o - F_c$ electron density map is contoured at 1.0 σ (blue mesh). The right subunit of Rv1816 is shown as orange ribbons. B, the isopropyl laurate binding site. Amino acid residues within 5.0 Å from the bound isopropyl laurate (green, carbon; red, oxygen) are included. The side chains of selected residues from the right subunit of Rv1816 in Fig. 7B are shown as orange sticks (orange, carbon; blue, nitrogen; red, oxygen). C, schematic representation of the Rv1816 and isopropyl laurate interactions. Amino acid residues within 5.0 Å from the bound isopropyl laurate are included.*

ligands of these proteins. The bound fatty acid ester is partially exposed to solvent at its isopropyl head-group, whereas the carbon chain tail of the molecule is buried in the region of the binding cavity created by α6, α8, and α9 and the loop region between the adjacent helices α9' and α10'. There are extensive hydrophobic interactions between the bound isopropyl laurate molecule and Rv1816. Within 5 Å of this bound ligand, at least 16 residues of Rv1816, including Pro-128, Thr-131, Ala-132, Ala-135, Thr-136, Val-139, Phe-142, Phe-143, Phe-177, Cys-189, Phe-190, Leu-192, Trp-193, Tyr- 208', Ala-210', and Met-212', are involved in the interaction (Fig. 9). Again, many of these residues are hydrophobic in nature.

Isothermal titration calorimetry

ITC was used to quantify regulator-ligand interactions. The strength of interaction between Rv3249c and the palmitate ligand was measured by ITC and possessed a negative enthalpic contribution of a typical hyperbolic binding curve (Fig. 10*A*). The enthalpic (ΔH) and entropic (ΔS) parameters of Rv3249c binding to palmitic acid are 629.6 ± 38.1 cal mol^{-1} and 21.3 cal mol^{-1} deg^{-1}, giving rise to a dissociation constant, K_D, of 7.5 ± 1.2 μM.
ITC was then employed to determine the binding affinity of the Rv1816 regulator for the laurate ligand (Fig. 10*B*). Because isopropyl laurate has a low solubility, we used lauric acid as a ligand for these experiments. Again, the titration depicts a typical hyperbolic binding curve, which is characterized by a negative enthalpic contribution. The thermodynamic parameters of binding of lauric acid to Rv1816 are 872.1 ± 41.4 cal mol^{-1} (ΔH) and 19.2 cal mol^{-1} deg^{-1} (ΔS), which give rise to a K_D of 14.4 ± 1.8 μM.

As the predicted binding sites of Rv3249c and Rv1816 within the promoter and intragenic region of the *mmpL* genes are partially overlapped, we thought that these two

regulators may share a similar set of ligands. We thus determined whether Rv1816 is able to bind palmitic acid (Fig. 10*C*). Surprisingly, ITC data depict that Rv1816 binds palmitic acid with a K_D of 23.2 ± 2.9 μM. The thermodynamic parameters of binding of palmitic acid to Rv1816 are 491.6 ± 29.2 cal mol^{-1} (Δ*H*) and 19.6 cal mol^{-1} deg^{-1} (Δ*S*).

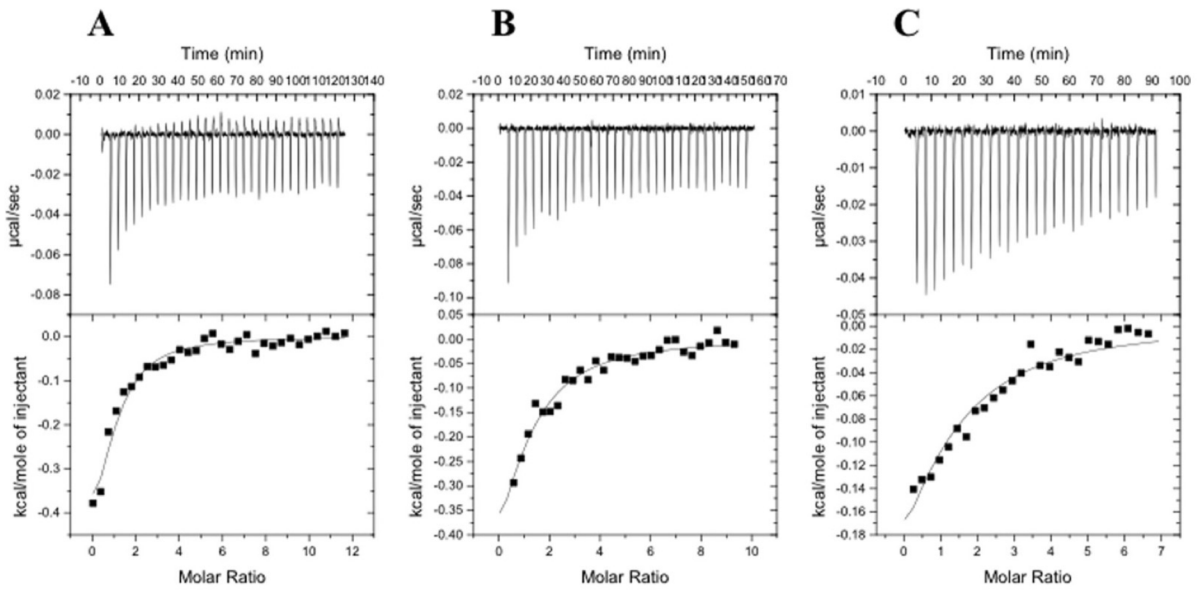

Figure 10. *Isothermal titration calorimetry for the binding of ligands. A, representative isothermal titration calorimetry for the binding of palmitic acid to Rv3249c. Upper panel, each peak corresponds to the injection of 10 μl of 500 μm palmitic acid in buffer containing 10 mm sodium phosphate (pH 7.2), 100 mm NaCl, and 0.001% DDM into the reaction containing 10 μm dimeric Rv3249c in the same buffer. Lower panel, cumulative heat of reaction is displayed as a function of the injection number. The solid line is the least square fit to the experimental data, giving a K_D of 7.5 ± 1.2 μm. B, representative isothermal titration calorimetry for the binding of lauric acid to Rv1816. Upper panel, each peak corresponds to the injection of 10 μl of 400 μm lauric acid in buffer containing 10 mm sodium phosphate (pH 7.2), 100 mm NaCl, and 0.001% DDM into the reaction containing 10 μm dimeric Rv1816 in the same buffer. Lower panel, cumulative heat of reaction is displayed as a function of the injection* number. *The solid line is the least square fit to the experimental data, giving a K_D of 14.4 ± 1.8 μm. C, representative isothermal titration calorimetry for the binding of palmitic acid to Rv1816. Upper panel, each peak corresponds to the injection of 10 μl of 400 μm palmitic acid in buffer containing 10 mm sodium phosphate (pH 7.2), 100 mm NaCl, and 0.001% DDM into the reaction containing 10 μm dimeric Rv1816 in the same buffer. Lower panel, cumulative heat of reaction is displayed as a function of the injection number. The solid line is the least square fit to the experimental data, giving a K_D of 23.2 ± 2.9 μm.*

Fluorescence polarization assay

Fluorescence polarization was used to quantify the strength of regulator DNA interactions. To identify regulatory targets of these proteins, we utilized chromatin immunoprecipitation sequencing (ChIP-Seq) data from Galagan and co-workers (26 –29) and the Tuberculosis Database. Regions of the *M. tuberculosis* H37Rv genome that were found by these experiments to interact with Rv3249c or Rv1816 were first examined to identify a potential binding sequence for each individual protein. Typically, TetR family proteins interact with DNAs via symmetric palindromic stretches called inverted repeats, 15–30 nucleotides long. Thus, the search was narrowed to include sequences that contain these patterns. For each protein we were able to identify a putative inverted repeat sequence located in one or more of the *M. tuberculosis* H37Rv genes encoding MmpL transporter proteins. These DNA sequences are in good agreement with both the consensus binding sequences and protein-DNA interactions determined by others (21). We found that the Rv3249c protein might act as a regulator for *mmpS1/L1*, *mmpL11*, and *rv1067c* and Rv1816 for *mmpL13b* and *rv1094*.

Fluorescence polarization assays were performed using purified regulator proteins and duplex DNA. For example, we quantified the interaction between Rv3249c and the 19-bp DNA sequence (ACATCGAAACGGTCGATGT), which is located at the *mmpL11* operon. Fig. 11*A* illustrates the binding isotherm of Rv3249c in the presence of 2.5 nM fluoresceinated DNA, indicating that Rv3249c binds this 19-bp promoter DNA with a dissociation constant, K_D, of 5.6 ± 1.0 nM. Similarly, the interaction between Rv3249c and the 19-bp putative promoter DNA sequence (ACCTCGCCGTAAACGATGT) within the *mmpS1/L1* operon was determined. The data suggest that the K_D value for this binding is

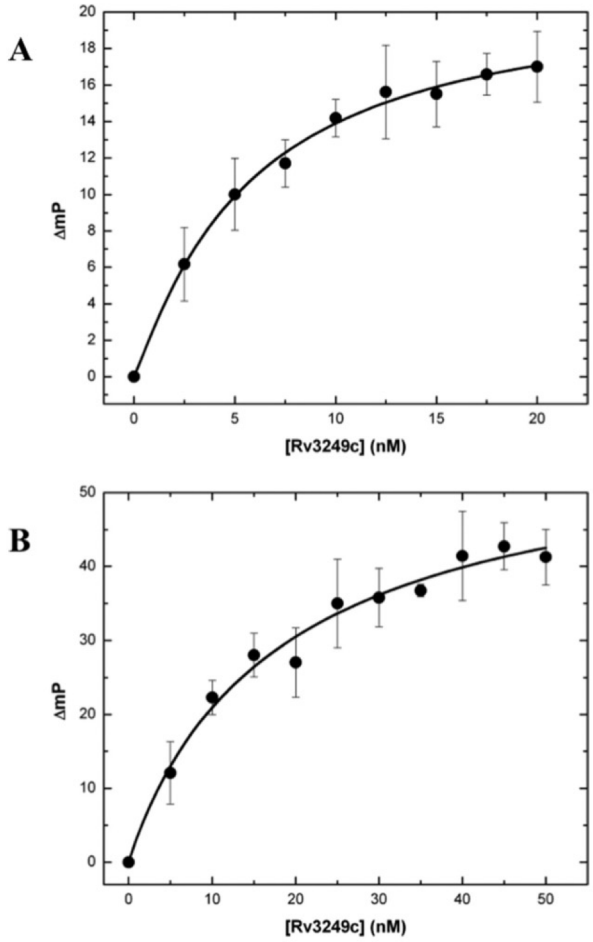

Figure 11. *Representative fluorescence polarization of Rv3249c. A, the binding isotherm of Rv3249c with the 27-bp DNA containing the 19-bp promoter sequence (ACATCGAAACGGTCGATGT) located at the mmpL11 operon, showing a K_D of 5.6 ± 1.0 nm. B, the binding isotherm of Rv3249c with the 29-bp DNA containing the 19-bp promoter sequence (ACCTCGCCGTAAACGATGT) within the mmpS1/L1 operon, showing a K_D of 18.7 ± 6.5 nm. Fluorescence polarization is defined by the equation FP = (V − H)/(V + H), where FP equals polarization, V equals the vertical component of the emitted light, and H equals the horizontal component of the emitted light of a fluorophore when excited by vertical plane polarized light. FP is a dimensionless entity and is not dependent on the intensity of the emitted light or on the concentration of the fluorophore. mP is related to FP, where 1 mP equals one-thousandth of FP.*

18.7 ± 6.5 nM (Fig. 11*B*). The binding constants of Rv3249c and Rv1816 with their corresponding DNAs are summarized in Table 2.

Electrophoretic mobility shift assay

Rv3249c is predicted to regulate expression of *mmpS1/L1, mmpS5/L5, mmpL10*, and *mmpL11* (Fig. 12*A*). We performed EMSAs using purified Rv3249c to demonstrate direct transcriptional regulation by Rv3249c. We observed a concentration-dependent shift of the *mmpL3*, *mmpL11*, and *mmpS1* probes (Fig. 12*B*). A second shift of the *mmpL3* probe with 0.5 μM Rv3249c suggests multiple binding sites for the regulator in this promoter region. As a control, EMSAs were performed in the presence of nonlabeled “cold”

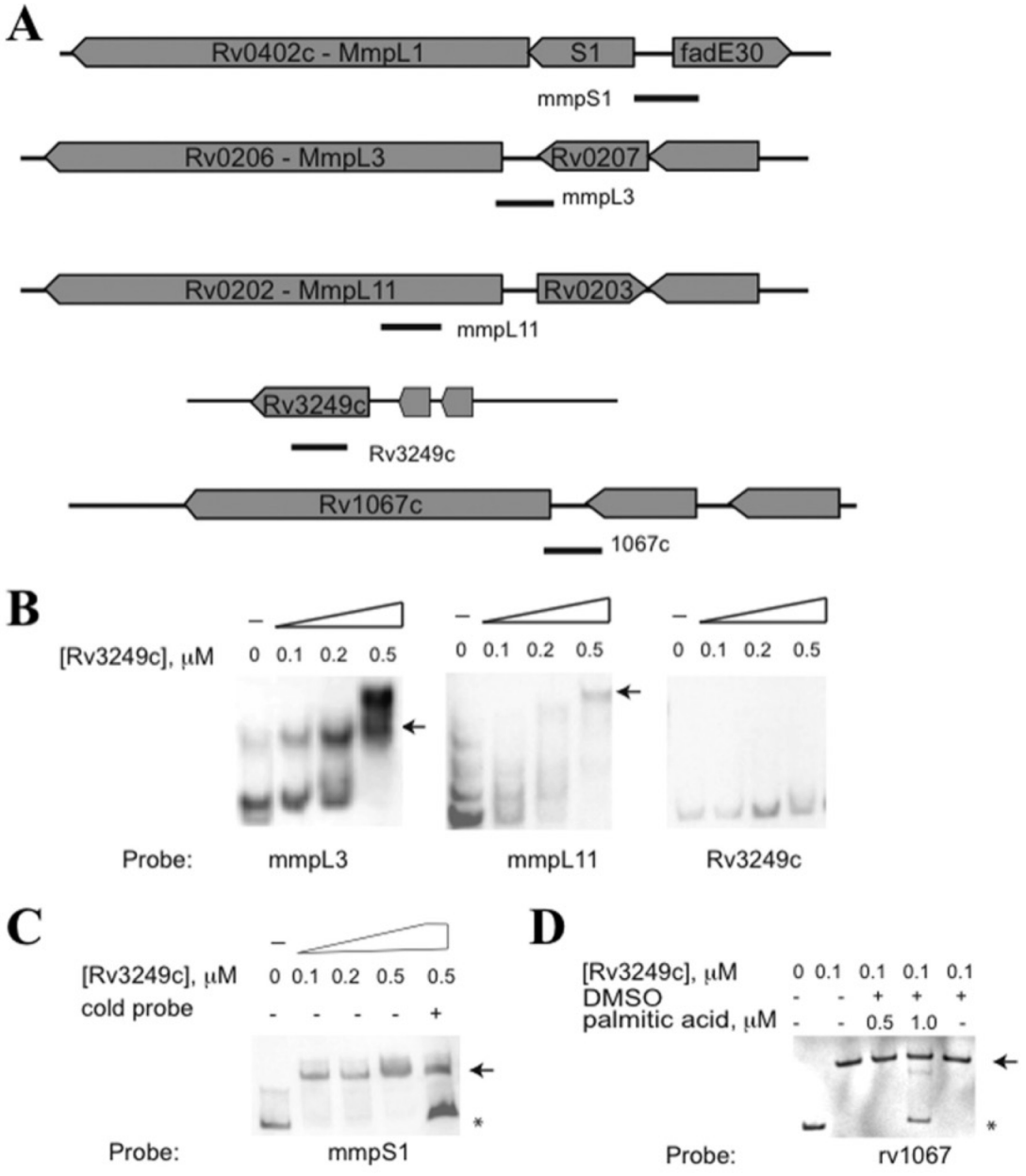

Figure 12. *Rv3249c binds to promoter regions of mmpS1 and mmpL3 and intragenic region of mmpL11. A, a schematic depicting the DNA probes used in EMSAs. B, for EMSA analysis, 6 nm Dig-labeled probe and the indicated micromolar concentrations of protein. C, to demonstrate specificity, the MmpS1 EMSA was performed in the presence of nonlabeled ("cold") probe. Reactions were performed with 6 nm Dig-labeled probe, the indicated micromolar concentrations of protein, and 360 nm cold probe. D, ligand-bound Rv3249c does not bind target probes. EMSA was performed using 12 nm Dig-labeled probe and 0.1 μm Rv3249c in the absence or presence of indicated concentration of palmitic acid. An arrow denotes the shifted probes and the asterisk notes the accumulation of free Dig-labeled probe.*

probe. Release of Dig-labeled probe was observed consistent with specific binding of Rv3249c to the *mmpS1* probe (Fig. 12*C*). The probe corresponding to a significant ChIP-Seq peak in the coding sequence of *rv3249c* itself did not shift. Our EMSA conditions may not be ideal for Rv3249c binding to the *rv3249c* probe or autoregulation may depend on a ligand or environmental condition that we have not adequately mimicked in our EMSA conditions. Rv3249c crystallized with a palmitic acid molecule, and ligand binding appears to be incompatible with DNA binding activity. We performed an EMSA in the presence and absence of palmitic acid to demonstrate this experimentally. The addition of ligand reduced binding of Rv3249c to the r*v1067c* probe (Fig. 12*D*). ChIP-Seq data suggest that Rv1816 potentially regulates itself, *mmpL2*, *mmpL3*, *mmpS4/L4*, *mmpL7*, *mmpL8*, and *mmpL11* expression. Interestingly ChIP-Seq data indicated that Rv1816 regulates *kasA*, which is co-regulated with *mmpL3* and encodes a β-ketoacyl-ACP synthase involved in meromycolate synthesis, in addition to the *mmpS3* gene, encoding an accessory MmpS protein. We performed EMSAs using purified Rv1816 and probes corresponding to Rv1816, the *mmpL3/mmpL11* region, *mmpL7*, *kasA*, and *mmpS3* (Fig. 13*A*). We observed a shift of the *mmpL3/mmpL11* locus probes, and less so for the *mmpL7* probe (Fig. 13*B*). We also obtained a robust concentration-dependent shift of the *kasA*, *mmpS3*, and *rv1094* probes. These data indicate that Rv1816 binds several times within the *mmpL3/mmpL11* genomic locus and suggest that Rv1816 contributes to the coordinated regulation of cell wall lipid biosynthesis and transport. As a control, EMSAs were performed in the presence of nonlabeled "cold" probe. Release of Dig-labeled probe was observed consistent with specific binding of Rv1816 to the *rv1094* probe (Fig. 13*C*). Rv1816 was co-crystallized with isopropyl laurate. In addition, ITC experiment suggested that this regulator binds lauric acid with an affinity in

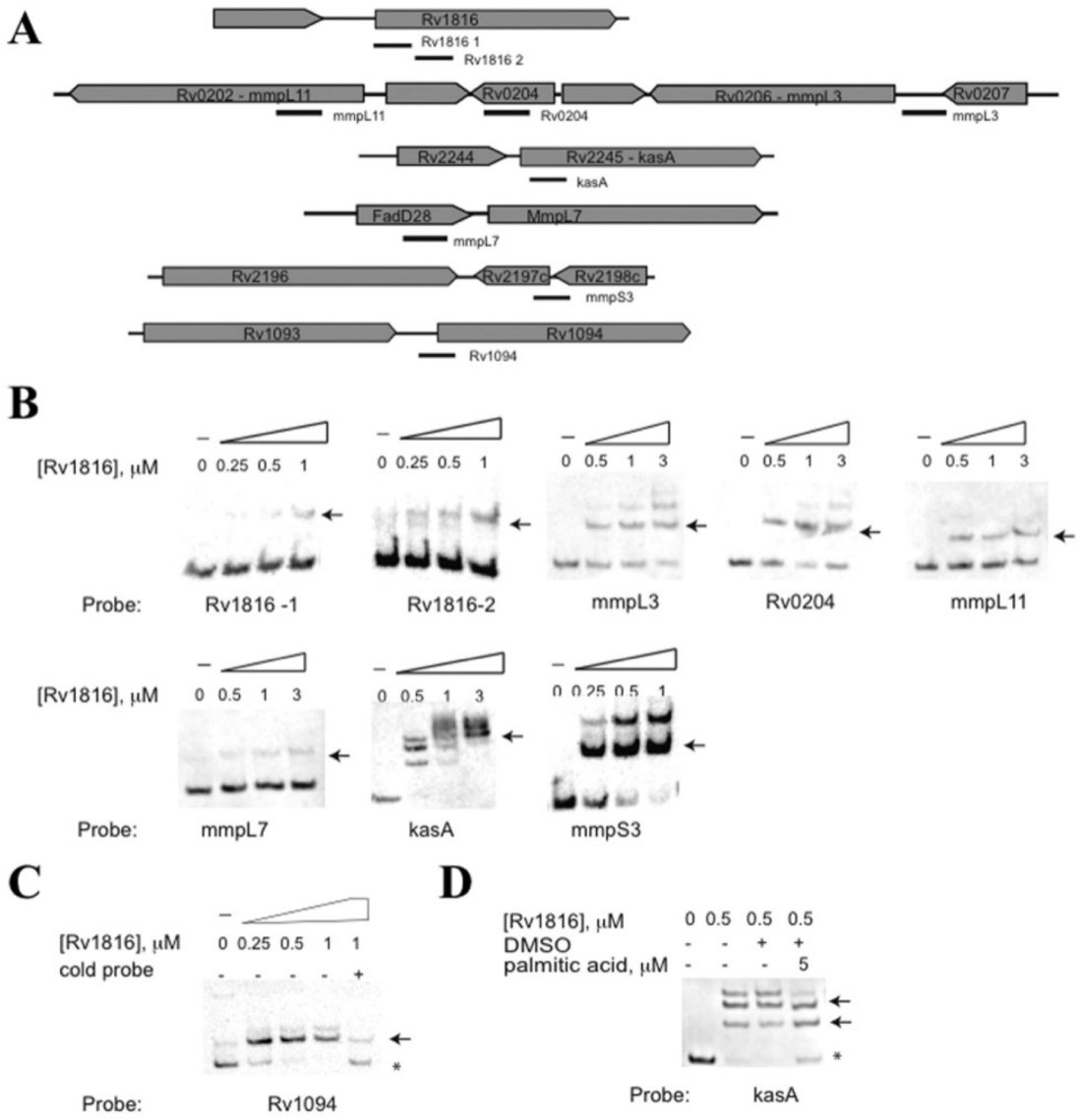

Figure 13. *Rv1816 binds to intragenic and/or promoter regions of rv1816, mmpL3, rv0204, mmpL11, mmpL7, kasA, and mmpS3. A, a schematic depicting the DNA probes used in EMSAs. B, EMSAs were performed using 12 nm Dig-labeled probe and the indicated micromolar concentrations of protein. An arrow denotes the shifted probes. EMSAs were performed using probes spanning regions of rv0550 and rv2245. These were predicted binding sites by ChIP-Seq but did not shift at the indicated concentrations. C, to demonstrate specificity, EMSAs were performed in the presence of nonlabeled ("cold") probe. Reactions were performed with 6 nm Dig-labeled probe, the indicated micromolar concentrations of protein, and 0.24 μm cold probe. The asterisk notes the accumulation of free Dig-labeled probe. D, ligand-bound Rv1816 does not bind target probes. EMSA was performed using 12 nm Dig-labeled probe and 0.5 μm Rv1816 in the absence or presence of 5 μm palmitic acid. An arrow denotes the shifted probes, and the asterisk notes the accumulation of free Dig-labeled probe.*

the micromolar range. We thus performed an EMSA in the presence of lauric acid. Surprisingly, the presence of this fatty acid did not reduce binding of Rv1816 to the *rv1094* probe. It appears that lauric acid binding does not prevent the interaction of this regulator with cognate DNAs. We then performed an EMSA both in the presence and absence of palmitic acid, because Rv1816 is also able to bind this compound. Indeed, the addition of ligand reduced binding of Rv1816 to the *rv1094* probe (Fig. 13*D*).

Conclusion

In this paper, we describe the crystal structures of the Rv3249c and Rv1816 transcriptional regulators, which participate in the regulatory network that controls the expression of essential and virulence-associated MmpL transporters. Specifically, existing ChIP-Seq data and the analyses presented herein suggest that Rv3249c regulates the genes *mmpS1/L1*, *mmpL3*, *mmpS5/L5*, *mmpL7*, *mmpL10*, *mmpL11*, and *mmpL12*, and Rv1816 regulates *mmpS2/L2*, *mmpL3*, *mmpS4/L4*, *mmpL7*, *mmpL8*, *mmpL11*, and *mmpL13b*. MmpL transporters significantly contribute to the export of important lipid components of the mycobacterial cell wall and are necessary for the virulence of this pathogen. Our experimental data demonstrate a direct binding of these transcriptional regulators to intragenic and promoter DNAs, providing evidence for the transcriptional control of *mmpL* gene expression. Multiple transcriptional factor binding sites exist within the promoter and intragenic region of the *mmpL* genes, and each transcriptional regulator recognizes several *mmpL* regulatory regions. These findings highlight that *mmpL* gene expression relies on a complex interplay of multiple transcription regulators.

Fortuitously, the crystal structures of Rv3249c and Rv1816 were resolved in

complexes with palmitic acid and isopropyl lauric acid, respectively. These structures suggest that saturated fatty acids are the natural substrates of these regulators. There are extensive interactions of these fatty acids with the transcriptional regulators, where the C-terminal regulatory domain of the regulators provides a hydrophobic environment for substrate binding. Within the fatty acid binding site of Rv3249c, there are at least 22 residues that participate in binding the palmitate ligand through hydrophobic interactions. Similarly, 16 residues are involved in binding the fatty acid ester isopropyl laurate within the ligand-binding pocket of Rv1816. Many of these residues are hydrophobic in nature, suggesting that fatty acid recognition is mainly governed by hydrophobic interactions.

The binding of palmitate to Rv3249c results in lengthening the center-to-center distance of this regulator, making it incompatible with cognate DNAs. However, the structure of Rv1816 indicates that binding of laurate does not change the center to center distance of the regulator. Thus, this regulator may still be able to hold onto the promoter DNA and repress gene expression even after interacting with this ligand. Indeed, gel shift experiment showed that the addition of laurate does not have any effect on the Rv1816-DNA complex. Rather, it appears that palmitic acid reduces DNA binding activity of Rv1816. Future work will define native fatty acid ligands for this regulator and dissect their effects on Rv1816 activity.

The TetR family regulators use several distinct mechanisms for modulating transcriptional regulation. However, the net consequence of binding of inducing ligands to these regulators is essentially the same. Ligand binding at the C-terminal regulatory domain triggers a long distance conformational change at the N-terminal DNA binding domain, resulting in the release of the regulator from its operator DNA. The TetR family regulators

utilize the N-terminal recognition helix α3 to bind the major groove of B-DNA. The distance between two consecutive major grooves of B-DNA is ~34 Å. One major mechanism found in the TetR family is that ligand binding increases the center to center distance between the two recognition helices α3 and α3' within the dimer, making this distance no longer compatible with the 34 Å separation between two successive major grooves of B-DNA. This results in the release of the dimeric regulator from the promoter region, allowing the expression of the respective regulated gene. For example, the center-to-center distance for QacR is 39 Å in the absence of inducing ligands (46). This distance becomes 48 Å upon ligand binding (47). The conformational change is augmented in the cases of CmeR (51) and EthR (50), where the center-to-center distances were observed to be 54 and 52 Å in the respective ligand-bound structures. Thus, it is most likely that ligand binding by the Rv3249c and Rv1816 regulators leading to derepression is primarily triggered by the increase in center-to-center distance, making these two regulators no longer compatible with their corresponding cognate DNAs.

References

1. Maartens, G., and Wilkinson R. J. (2007) Tuberculosis. *Lancet* **370,** 2030 –2043

2. Frieden, T. R., Sterling, T., Pablos-Mendez, A., Kilburn, J. O., Cauthen, G. M., and Dooley, S. W. (1993) The emergence of drug-resistant tuberculosis in New York City. *N. Engl. J. Med.* **328,** 521–556

3. Pillay, M., and Sturm, A. W. (2007) Evolution of the extensively drug- resistant F15/LAM4/KZN strain of *Mycobacterium tuberculosis* in Kwa-Zulu-Natal, South Africa. *Clin. Infect. Dis.* **45,** 1409 –1414

4. Goldman, R. C., Plumley, K. V., and Laughon, B. E. (2007) The evolution of extensively drug resistant tuberculosis (XDR–TB): history, status and issues for global control. *Infect. Disord. Drug Targets* **7,** 73–91

5. Udwadia, Z. F., Amale, R. A., Ajbani, K. K., and Rodrigues, C. (2012) Totally drug-resistant tuberculosis in India. *Clin. Infect. Dis.* **54,** 579 –581

6. Iseman M. D. (1993) Treatment of multidrug-resistant tuberculosis. *N. Engl. J. Med.* **329,** 784 –791

7. Domenech, P., Reed, M. B., and Barry, C. E., 3rd (2005) Contribution of the *Mycobacterium tuberculosis* MmpL protein family to virulence and drug resistance. *Infect. Immun.* **73,** 3492–3501

8. Jain, M., and Cox, J. S. (2005) Interaction between polyketide synthase and transporter suggests coupled synthesis and export of virulence lipid in *M. tuberculosis*. *PLoS Pathog*. **1,** 12–19

9. Cox, J. S., Chen, B., McNeil, M., and Jacobs, W. R., Jr. (1999) Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* **402,** 79 – 83

10. Brennan, P. J., and Nikaido, H. (1995) The envelope of mycobacteria. *Annu. Rev. Biochem.* **64,** 29 – 63

11. Converse, S. E., Mougous, J. D., Leavell, M. D., Leary, J. A., Bertozzi, C. R., and Cox, J. S. (2003) MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 6121– 6126

12. Marrakchi, H., Lanéelle, G., and Quémard, A. (2000) InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II. *Microbiology* **146,** 289 –296

13. Luckner, S. R., Machutta, C. A., Tonge, P. J., and Kisker, C. (2009) Crystal structures of *Mycobacterium tuberculosis* KasA show mode of action within cell wall biosynthesis and its inhibition by thiolactomycin. *Structure* **17,** 1004 –1013

14. Winder, F. G., and Collins, P. B. (1970) Inhibition by isoniazid of synthesis of mycolic acids in *Mycobacterium tuberculosis*. *Microbiology* **63,** 41– 48

15. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393,** 537–544

16. Tseng, T. T., Gratwick, K. S., Kollman, J., Park, D., Nies, D. H., Goffeau, A., and Saier, M. H., Jr. (1999) The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol.* **1,** 107–125

17. Tekaia, F., Gordon, S.V., Garnier, T., Brosch, R., Barrell, B.G., and Cole, S. T. (1999) Analysis of the proteome of *Mycobacterium tuberculosis* in silico. *Tuber. Lung. Dis.* **79,** 329 –342

18. Milano, A., Pasca, M. R., Provvedi, R., Lucarelli, A. P., Manina, G., Ribeiro, A. L., Manganelli, R., and Riccardi, G. (2009) Azole resistance in *Mycobacterium tuberculosis* is mediated by the MmpS5-MmpL5 efflux system. *Tuberculosis* **89,** 84 – 90

19. Wells, R. M., Jones, C. M., Xi, Z., Speer, A., Danilchanka, O., Doornbos, K. S., Sun, P., Wu, F., Tian, C., and Niederweis, M. (2013) Discovery of a siderophore export system essential for virulence of *Mycobacterium tuberculosis*. *PLoS Pathog.* **9,** 1–14

20. Pacheco, S.A., Hsu, F.-F., Powers, K.M., and Purdy, G.E. (2013) MmpL11 transports mycolic acid-containing lipids to the mycobacterial cell wall and contributes to biofilm formation in *M. smegmatis*. *J. Biol. Chem.* **288,** 24213–24222

21. Minch, K.J., Rustad, T.R., Peterson, E.J., Winkler, J., Reiss, D.J., Ma, S., Hickey, M., Brabant, W., Morrison, B., Turkarslan, S., Mawhinney, C., Galagan, J. E., Price, N. D., Baliga, N. S., and Sherman, D. R. (2015) The DNA-binding network of *Mycobacterium tuberculosis*. *Nat. Commun.* **6,** 5829

22. Varela, C., Rittmann, D., Singh, A., Krumbach, K., Bhatt, K., Eggeling, L., Besra, G. S., and Bhatt, A. (2012) MmpL genes are associated with mycolic acid metabolism in mycobacteria and corynebacteria. *Chem. Biol.* **19,** 498 –506

23. Grzegorzewicz, A. E., Pham, H., Gundi, V. A., Scherman, M. S., North, E. J., Hess, T., Jones, V., Gruppo, V., Born, S. E., Korduláková, J., Chavadi, S. S., Morisseau, C., Lenaerts, A. J., Lee, R. E., McNeil, M. R., and Jackson, M. (2012) Inhibition of mycolic acid transport across the *Mycobacterium tuberculosis* plasma membrane. *Nat. Chem. Biol.* **8,** 334 –341

24. Camacho, L. R., Constant, P., Raynaud, C., Laneelle, M. A., Triccas, J. A., Gicquel, B., Daffe, M., and Guilhot, C. (2001) Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*: evidence that this lipid is involved in the cell wall permeability barrier. *J. Biol. Chem.* **276,** 19845–19854

25. Belardinelli, J. M., Larrouy-Maumus, G., Jones, V., Sorio de Carvalho, L. P., McNeil, M. R., and Jackson, M. (2014) Biosynthesis and translocation of unsulfated acyltrehaloses in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **289,** 27952–27965

26. Reddy, T. B., Riley, R., Wymore, F., Montgomery, P., DeCaprio, D., Engels, R., Gellesch, M., Hubble, J., Jen, D., Jin H, Koehrsen, M., Larson, L., Mao, M., Nitzberg, M., Sisk, P., Stolte, C., Weiner, B., White, J., Zachariah, Z. K., Sherlock, G., Galagan, J. E., Ball, C. A., and Schoolnik, G. K. (2009) TB database: an integrated platform for tuberculosis research. *Nucleic Acids Res.* **37,** D499 –D508

27. Galagan, J. E., Sisk, P., Stolte, C., Weiner, B., Koehrsen, M., Wymore, F., Reddy, T. B., Zucker, J. D., Engels, R., Gellesch, M., Hubble, J., Jin, H., Larson, L., Mao, M., Nitzberg, M., White, J., Zachariah, Z. K., Sherlock, G., Ball, C. A., and Schoolnik, G. K. (2010) TB database 2010: overview and update. *Tuberculosis* **90,** 225–235

28. Rustad, T. R., Minch, K. J., Ma, S., Winkler, J. K., Hobbs, S., Hickey, M., Brabant, W., Turkarslan, S., Price, N. D., Baliga, N. S., and Sherman, D. R. (2014) Mapping and manipulating the *Mycobacterium tuberculosis* transcriptome using a transcription factor overexpression-derived regulatory network. *Genome Biol.* **15,** 502

29. Galagan, J. E., Minch, K., Peterson, M., Lyubetskaya, A., Azizi, E., Sweet, L., Gomes, A., Rustad, T., Dolganov, G., Glotova, I., Abeel, T., Mahwinney, C., Kennedy, A. D., Allard, R., Brabant, W., Krueger, A., Jaini, S., Honda, B., Yu, W. H., Hickey, M. J., Zucker, J., Garay, C., Weiner, B., Sisk, P., Stolte, C., Winkler, J. K., Van de Peer, Y., Iazzetti, P., Camacho, D., Dreyfuss, J., Liu, Y., Dorhoi, A., Mollenkopf, H. J., Drogaris, P., Lamontagne, J., Zhou, Y., Piquenot, J., Park, S. T., Raman, S., Kaufmann, S. H., Mohney, R. P., Chelsky, D., Moody, D. B., Sherman, D. R., and Schoolnik, G. K. (2013) The *Mycobacterium tuberculosis* regulatory network and hypoxia. *Nature* **499,** 178 –183

30. Radhakrishnan, A., Kumar, N., Wright, C. C., Chou, T.-H., Bolla, J. R., Tringides, M. L., Rajashankar, K. R., Su, C.-C., Purdy G. E., and Yu, E. W. (2014) Crystal structure of the transcriptional regulator Rv0678 of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **289,** 16526 –16540

31. Grkovic, S., Brown, M. H., and Skurray, R. A. (2002) Regulation of bacterial drug export systems. *Microbiol. Mol. Biol. Rev.* **66,** 671–701

32. Ramos, J. L., Martínez-Bueno, M., Molina-Henares, A. J., Terán, W., Watanabe, K., Zhang, X., Gallegos, M. T., Brennan, R., and Tobes, R. (2005) The TetR family of transcriptional repressors. *Microbiol. Mol. Biol. Rev.* **69,** 326 –356

33. Chou, T., Delmar, J. A., Wright, C. C., Kumar, N., Radhakrishnan, A., Doh, J. K., Licon, M. H., Bolla, J. R., Lei, H., Rajashankar, K. R., Su, C.-C., Purdy G. E., and Yu, E. W. (2015) Crystal structure of the *Mycobacterium tuberculosis* transcriptional regulator Rv0302. *Protein Sci.* 10.1002/pro.2802

34. Otwinowski, Z., and Minor, M. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326

35. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1772–1779

36. Pape, T., and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* **37,** 843– 844

37. Otwinowski, Z. (1991) ML-PHARE, CCP4 Proc. 80 – 88, Daresbury Laboratory, Warrington, UK

38. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760 –763

39. Cowtan, K. (2010) Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 470 – 478

40. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126 –2132

41. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., Mc- Coy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1948 –1954

42. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921

43. Terwilliger, T. C. (2001) Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallogr. D Biol. Crystallogr.* **57,** 1755–1762

44. Hinrichs, W., Kisker, C., Düvel, M., Müller, A., Tovar, K., Hillen, W., and Saenger, W. (1994) Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. *Science* **264,** 418 – 420

45. Orth, P., Schnappinger, D., Hillen, W., Saenger, W., and Hinrichs, W. (2000) Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat. Struct. Biol.* **7,** 215–219

46. Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A., and Brennan, R. G. (2001) Structural mechanisms of QacR induction and multidrug recognition. *Science* **294,** 2158 –2163

47. Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A., and Brennan, R. G. (2002) Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J.* **21,** 1210 –1218

48. Natsume, R., Ohnishi, Y., Senda, T., and Horinouchi, S. (2004) Crystal structure of a γ-butyrolactone autoregulator receptor protein in *Streptomyces coelicolor* A3$_2$. *J. Mol. Biol.* **336,** 409 – 419

49. Dover L. G., Corsino, P. E., Daniels, I. R., Cocklin, S. L., Tatituri, V., Besra, G. S., and Fütterer, K. (2004) Crystal structure of the TetR/CamR family repressor *Mycobacterium tuberculosis* EthR implicated in ethionamide resistance. *J. Mol. Biol.* **340,** 1095–1105

50. Frénois F, Engohang-Ndong, J., Locht, C., Baulard, A. R., and Villeret, V. (2004) Structure of EthR in a ligand bound conformation reveals therapeutic perspectives against tuberculosis. *Mol. Cell* **16,** 301–307

51. Gu, R., Su, C.-C., Shi, F., Li, M., McDermott, G., Zhang, Q., and Yu, E. W. (2007) Crystal structure of the transcriptional regulator CmeR from *Campylobacter jejuni*. *J. Mol. Biol.* **372,** 583–593

52. Lei, H. T., Shen, Z., Surana, P., Routh, M. D., Su, C.-C., Zhang, Q., Yu, E. W. (2011) Crystal structures of CmeR-bile acid complexes from *Campylobacter jejuni*. *Protein Sci.* **20,** 712–723

53. Li, M., Gu, R., Su, C.-C., Routh, M. D., Harris, K. C., Jewell, E. S., McDermott, G., Yu, E. W. (2007) Crystal structure of the transcriptional regulator AcrR from *Escherichia coli*. *J. Mol. Biol.* **374,** 591– 603

54. Hernández, A., Maté, M. J., Sánchez-Díaz, P. C., Romero, A., Rojo, F., and Martínez, J. L. (2009) Structural and functional analysis of SmeT, the repressor of the *Stenotrophomonas maltophilia* multidrug efflux pump SmeDEF. *J. Biol. Chem.* **284,** 14428 –14438

55. Bolla, J. R., Do, S. V., Long, F., Dai, L., Su, C.-C., Lei, H.-T., Chen, X., Gerkey, J. E., Murphy, D. C., Rajashankar, K. R., Zhang, Q., and Yu, E. W. (2012) Structural and functional analysis of the transcriptional regulator Rv3066 of *Mycobacterium tuberculosis*. *Nucleic Acids Res.* **40,** 9340 –9355

56. Kumar, N., Radhakrishnan, A., Wright, C. C., Chou, T.-H., Lei, H.-T., Bolla, J. R., Tringides, M. L., Rajashankar, K. R., Su, C.-C., Purdy, G. E., and Yu, E. W. (2014) Crystal structure of the transcriptional regulator Rv1219c of *Mycobacterium tuberculosis*. *Protein Sci.* **23,** 423– 432

CHAPTER IV

ASSEMBLY OF THE NEISSERIA GONORRHOEAE MTRCDE TRIPARTITE MULTIDRUG EFFLUX SYSTEM

Introduction

Neisseria gonorrhoeae is a Gram-negative human pathogen and the cause of the sexually transmitted disease gonorrhea. Recently, treatment of gonorrhea has been severely compromised by the development of antibiotic-resistance strains. For example, the use of ciprofloxacin, formerly a first-line antibiotic, has been discontinued since 2007, when nearly 14% of all *N. gonorrhoeae* infections in the US were found to exhibit some resistance to this drug (https://www.cdc.gov/std/gonorrhea/arg/basic.htm). Among antimicrobial-resistance mechanisms, active transport is now recognized as the major means by which pathogens survive antibiotic regimens.[1,2] To remove toxic molecules from the cell, bacteria utilize powerful pumping machinery, composed of one or more protein components. These multidrug efflux pumps are capable of actively pumping out a broad range of noxious agents directly from the bacterial cell.[3] Metal-ion and antibiotic resistance systems of this type have been found encoded on plasmids of every bacterial species tested.

The best characterized efflux system in *N. gonorrhoeae* is the MtrCDE multidrug efflux system. The tripartite complex, which spans both the inner and outer membranes, is formed by interactions between the inner membrane RND transporter MtrD, the periplasmic membrane fusion protein MtrC, and the outer membrane channel MtrE (Fig. 1). This powerful efflux pump mediates the export of hydrophobic antimicrobial agents, such as antibiotics, nonionic detergents, certain antibacterial peptides, bile salts, and gonadal steroidal hormones.[4-9]

Recently, our group has determined the crystal structure of the MtrE outer membrane channel, revealing the first open conformation in this protein family.[8] We have also determined the crystal structure of the inner membrane RND transporter MtrD.[10] The availability of the three-dimensional structures of these efflux transporters and their accessory proteins should allow us to block their function; potentially by the rational design of inhibitors, as demonstrated by AcrB and DARPin.[11,12] However, the direct approach, to interrupt the assembly of the tripartite complex, necessitates a high-resolution of a completely assembled RND tripartite efflux complex. In the following chapter, we discuss our strategies to elucidate the crystal structure of the MtrCDE pump, which have resulted in a preliminary model of the bipartite MtrCD complex to 5.5 Å resolution.

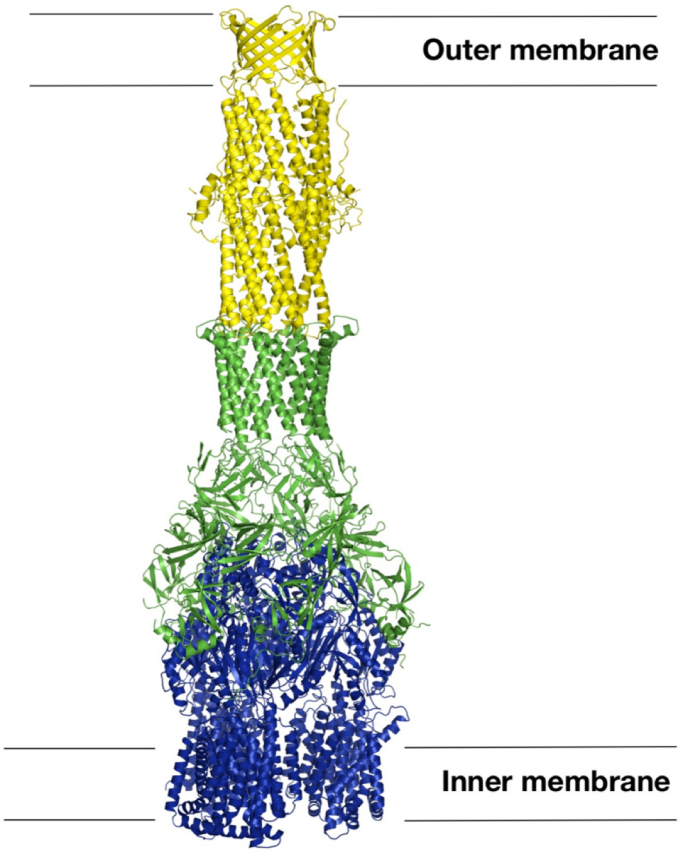

Figure 1. *Ribbon diagram model of the MtrCDE tripartite efflux system. Each protein is labeled with a different color (yellow, MtrE; green, MtrC; blue, MtrD). The figure was prepared using PyMOL.*

Co-crystal Structure of MtrCD

Initial screening of the bipartite MtrCD complex was carried out using hanging-drop vapor diffusion, with a 2:1 mixture of MtrCΔNT (MtrC lacking the N-terminal 34 amino acids) and the MtrD construct used to obtain our previous crystal structure.[10] Using a well solution of 33% Polyethylene glycol (PEG 400), 0.1 M $MgCl_2$, 0.1 M NaCl, and 0.1M Tris (pH 8.5), with the additions of 1% Sucrose monodecanoate (SM) and 0.1 mM Carbonyl cyanide m-chlorophenyl hydrazone (CCCP), small, approximately 50 μm, diamond-shaped crystals grew from the cover slip after approximately 3 weeks. The preliminary structure of the MtrCD bipartite complex was determined to a resolution of 5.5 Å, using molecular replacement with our previous structure of MtrD (PDB ID: 4MT1) (Fig. 2). This MtrCD cocrystal reveals a stoichiometry (3:6, MtrD:MtrC) consistent with previous experiments

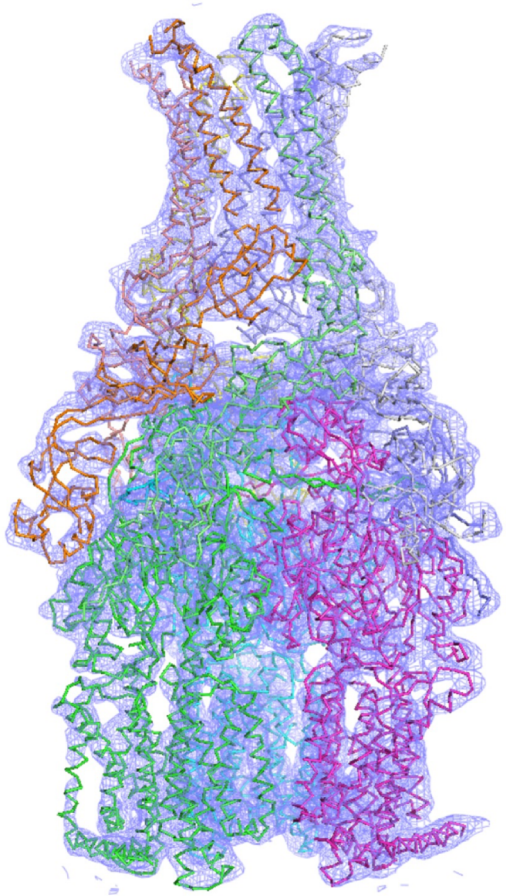

Figure 2. *Experimental electron density map of the MtrCD bipartite compelx at a resolution of 5.50 Å. The electron density maps are contoured at 1.0 σ. The Cα traces of the three MtrD molecules in the trimer are colored green, cyan, and magenta. The Cα traces of the six MtrC molecules in the hexamer are colored green, lavender, white, light pink, orange, and yellow.*

for the Mtr system[13] and recent cryo-electron microscopy (cryo-EM) structures of the AcrAB-TolC[14-16] and MexAB-OprM[16] efflux pumps.

Unfortunately, after extensive crystallization and detergent screening, we were not able to improve upon this model. Further, single particle analysis by cryo-EM did not reveal the presence of a significant number of intact complexes in our purified protein samples. Thus, a major effort was dedicated to improving the affinity of these proteins for a stable tripartite efflux complex.

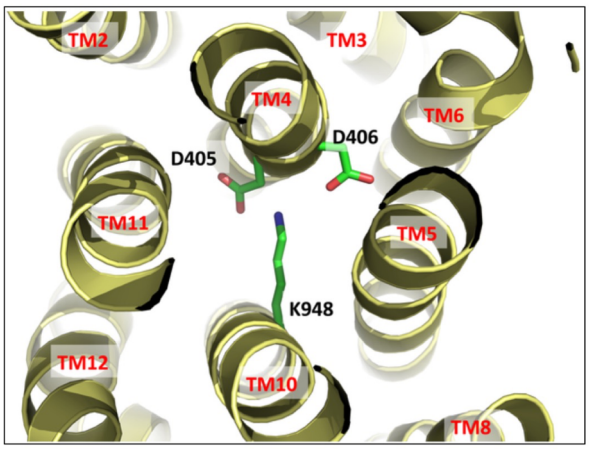

Figure 3. *Ion pairs of MtrD in the transmembrane domain viewed from the cytoplasmic side. Residues D405 and D406 of TM4 and K948 of TM10 that form ion pairs, which may play an important role in proton translocation, are in green sticks. Figure and caption borrowed from Bolla et al.*[10]

MtrD Proton-Motive Force Mutations

Efflux by RND transporters is dependent on the proton motive force (PMF). Previous studies of AcrB and MexB revealed three highly conserved residues (D407, D408, and K939) that are essential to the pump function and the proposed PMF pathway.[17-20] Based on sequence alignment, the corresponding charged residues D405 and D406 of TM4 and K948 of TM10 are necessary for the PMF of MtrD (Fig. 3). Janganan et al. found that mutation of these residues to lysine (MtrD D405K or D406K) or glutamic acid (K948E), rendered cells expressing the mutant pumps more susceptible to the antibiotics tetracycline, erythromycin, and novobiocin. Further, the authors concluded that it is possible that MtrE is unable to dissociate from the MtrCD D405K/D406K/E948E complex.[17] Thus, these

mutations are ideal for assembly of the tripartite complex.

Although the single point mutations MtrD D405K, D406K, or E948E, as well as the double mutation D405K D406K, severely attenuated the expression of the pump, we were able to purify ~1 mg of each protein using similar procedures as for wild-type MtrD.

Initial crystal screening of the bipartite complex was done by hanging-drop vapor diffusion, using a 2:1 mixture of MtrC:MtrD D405K D406K and 1% sucrose monodecanoate (SM) as the protein solution. Crystals were obtained for well solutions of: i) 6.6% PEG 4000, 0.1 M NaCl, 0.02 M K-MES (pH 6.7); and ii) 28% PEG 400, 0.1 M Lithium citrate, 0.1 M Glycine (pH 9.3). Unfortunately, these conditions could not be improved to obtain high-quality diffraction data.

MtrC Crosslinking

Janganan et. al also reported success in stabilizing high-order complexes of both MtrC and the MtrCE bipartite complex using a chemical crosslinking strategy.[13] The chemical crosslinker bismaleimidohexane (BMH) forms a covalent ~13 Å bond between sulfhydryl groups of neighboring Cysteine residues. As the MtrCΔNT construct lacks any Cysteine residues, the single point mutation E149C ensures that only two neighboring MtrCΔNT molecules will be bonded into a chemically stabilized dimer. Using our low-resolution bipartite cocrystal structure as a model, these residues were found to be 14.9 Å apart in the MtrC hexamer (Fig. 4). As this strategy could be used to stabilize the MtrC hexamer, and MtrCE bipartite complex, it is also possible that the stable tripartite complex could be formed *via* crosslinking of MtrC.

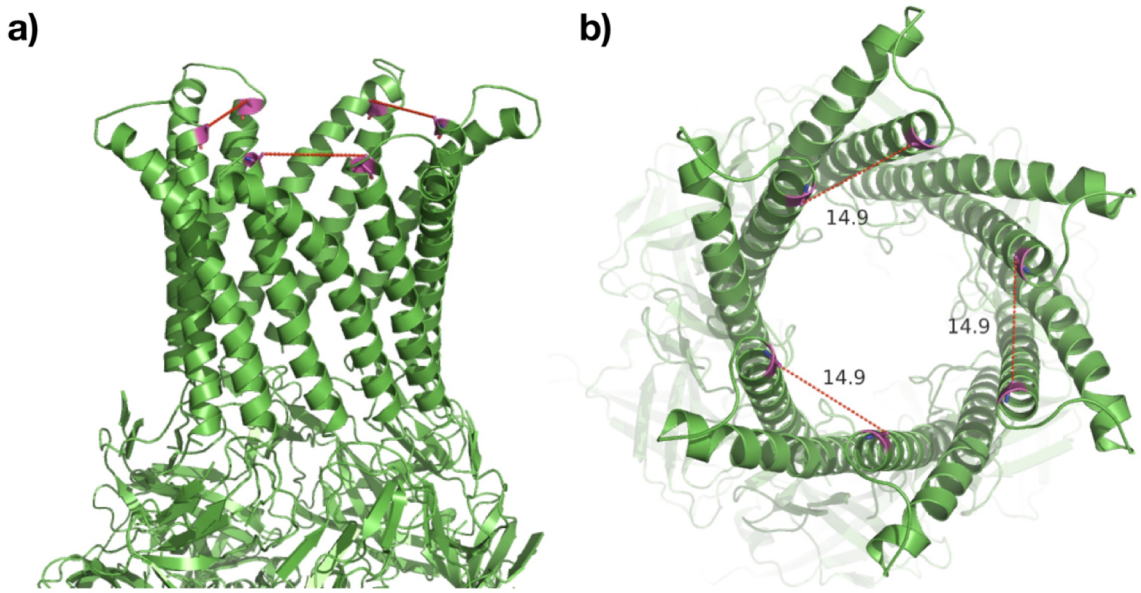

Figure 4. *Ribbon diagram of the hairpin domain of the MtrC hexamer as viewed from a) the periplasmic side and b) top-down, from the inner-membrane. Each molecule is colored in green and the residue E149, used for cross-linking, is highlighted in magenta. The distance between adjacent E149 residues is depicted in red dashed lines and calculated using PyMol.*

Thus, we purified MtrCΔNT E149C (Fig. 5a) and subjected it to crosslinking by BMH. After incubation with BMH at room temperature for 1 hr, a strong upper band appeared in SDS-PAGE analysis of the protein sample, corresponding to the molecular weight of the MtrCΔNT dimer (Fig. 5b). Additional higher molecular weight bands also appeared in the sample, presumably corresponding to higher order complexes of MtrCΔNT E149C. These dimeric and hexameric bands were not present in the sample before BMH crosslinking.

Size exclusion chromatography (SEC) analysis of the cross-linked sample resulted in a single peak, which eluted at a much lower volume (much higher molecular weight) compared to MtrCΔNT alone. When this sample was analyzed by SDS-PAGE, three prominent bands were visible, which correspond very closely to the theoretical monomeric,

dimeric, and hexameric weights of MtrC (Fig. 6). Indeed, cross-linking of MtrC by BMH does appear to stabilize the expected native hexamer.

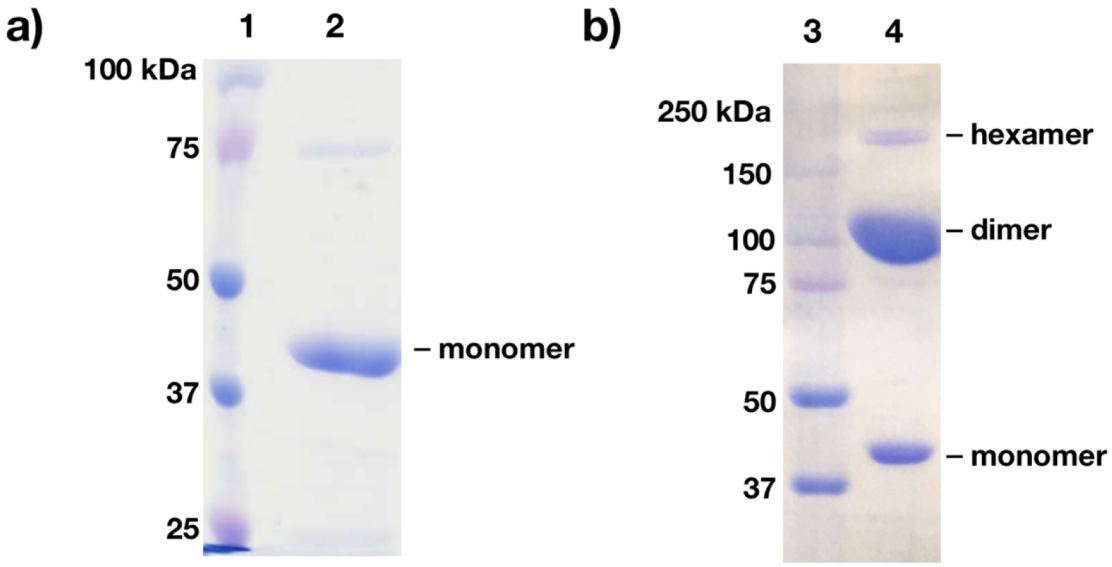

Figure 5. *SDS-PAGE analysis of purified MtrCΔNT E149C a) before and b) after cross-linking with BMH. a) Lane 1, protein molecular weight standards. Lane 2, MtrCΔNT E149C. b) Lane 3, protein molecular weight standards. Lane 4, 2:1, MtrCΔNT E149C:BMH.*

When this sample was mixed with MtrD in a 2:1 MtrC:MtrD ratio and subjected to crystal screening by hanging drop vapor diffusion, easily reproducible crystals were obtained from a well solution of 12% PEG 4000, 0.1 M $(NH_4)_2SO_4$ or NH_4-citrate dibasic or NH_4-citrate tribasic, 0.1 M Na-citrate (pH 5.6), and 22% glycerol. Unfortunately, we were not able to improve upon this condition and high-resolution data is unavailable. We have also attempted to combine this approach with the previous MtrD PMF mutants. A 2:1 mixture of MtrCΔNT E149C with the single-point mutants MtrD D405A, D406A, or the double mutant D405A D406A, respectively, did not result in any crystals which diffracted to high resolution. It is interesting to note that by mutation to alanine, instead of lysine, expression of these MtrD mutants greatly improved.

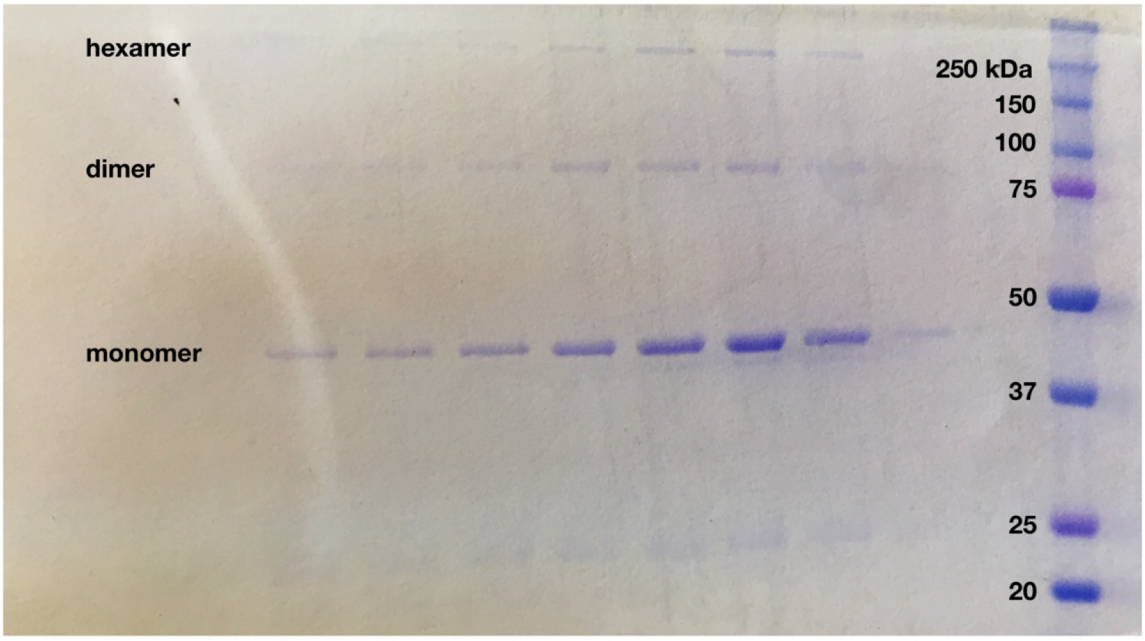

Figure 6. *After crosslinking, the MtrCΔNT E149C protein was passed through a Superdex 200 gel filtration column. A single peak was collected and analyzed by SDS-PAGE, revealing monomeric, dimeric and hexameric species.*

Chimeric Protein Construction

Recently, one of the first low-resolution models of a tripartite efflux pump was obtained in the AcrABZ-TolC system.[14] This was achieved, in part, by the use of chimeric protein constructs that fused one AcrA molecule to a loop region preceding the first large periplasmic domain of AcrB. This construct was found to preserve the periplasmic location of AcrA and the native interaction between AcrA and AcrB, while stabilizing this complex by forcing AcrA and AcrB in close proximity.

Similarly, we designed chimeric constructs of MtrC and MtrD. Utilizing our co-crystal structure, we located a flexible loop region immediately preceding TM helix α2 of MtrD. At residue P324 of MtrD, we inserted one molecule of MtrCΔNT, flanked by two

flexible 6xGS linkers at the N- and C-termini, to allow MtrC and MtrD to adopt their native conformation (Fig. 7). Preceding TM helix α8, and after the second large periplasmic domain of MtrD, is a similar flexible region. We designed a second construct containing one molecule of MtrCΔNT, with flexible N- and C-termini, in this region that precedes MtrD residue G858. While we were able to obtain a small signal, corresponding to expression of these MtrCD chimeric proteins in *E. coli* C43(DE3), (Fig. 8) we were not able to optimize the process to produce the mg quantities of protein needed for crystallization trials.

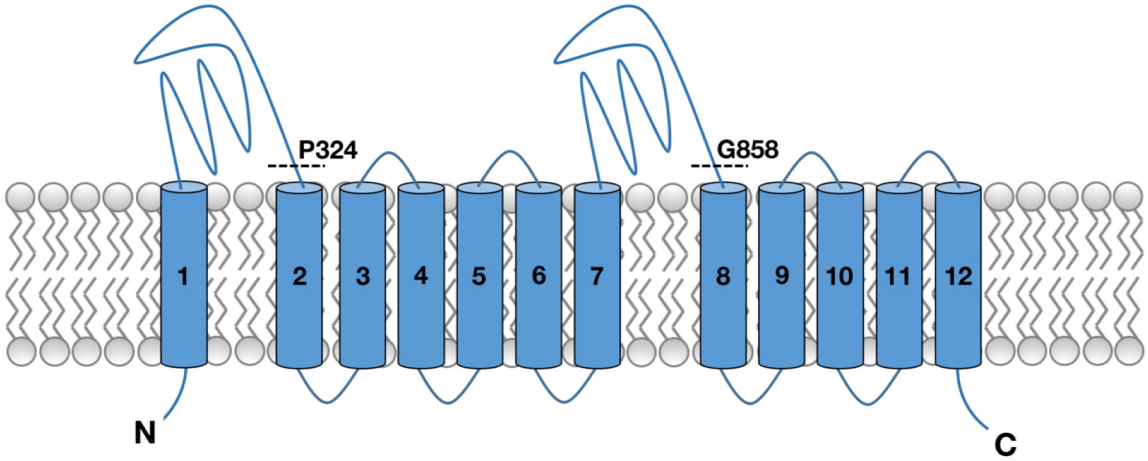

Figure 7. *Topology of the MtrD protein. View from the side of the inner-membrane. To create the MtrCD chimera, we inserted one MtrC molecule flanked by flexible linkers preceding TM helix α2 at residue P324 or preceding TM helix α8 at residue G858 (dotted line).*

MtrC Engineering

The essential link in tripartite RND efflux complexes is the periplasmic membrane fusion protein. In the highest resolution structures of these complexes that are currently available, there is no evidence of interaction between the inner-membrane transporter and the outer-membrane channel.[14,16] Instead, the periplasmic adaptor protein alone mediates the

interaction between these components, forming a bridge between inner and outer membranes to expel diverse antimicrobials from the cytoplasm all the way to the cell exterior.

Based on the currently available structural information, these proteins can be divided into 5 domains. The three β-sheet domains, comprising the lipoyl, β-barrel, and membrane proximal domains, primarily interact with the two large periplasmic domains of the RND transporter, while the α-helical hairpin domain is the only domain to contact the outer membrane channel (Fig. 9) Finally, the N-terminal residues of these proteins typically contain both a signal peptide (SP) sequence, targeting the protein to the periplasm, as well as an anchoring sequence for permanent attachment to the periplasmic side of the inner-membrane.[21-30]

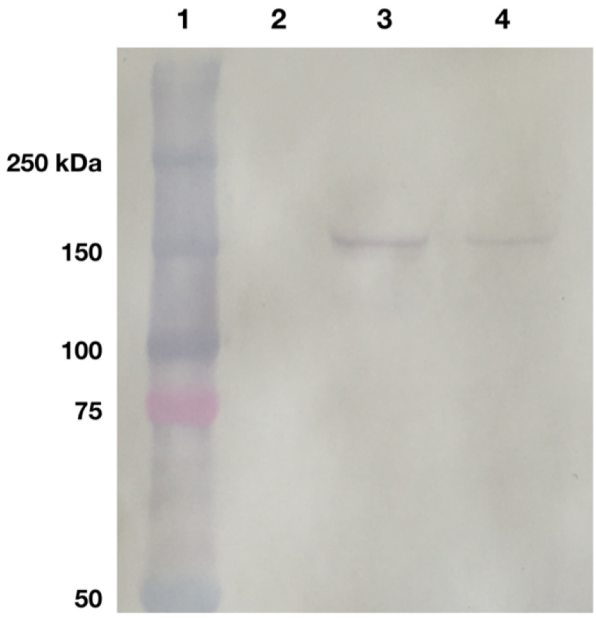

Figure 8. *Western analysis of the expression of the MtrCD chimeric proteins. Lane 1, protein molecular weight standards. Lane 2, empty. Lanes 3 and 4, MtrCD chimeric proteins.*

The mechanisms of inner-membrane anchoring among the membrane fusion protein family are diverse. For example, AcrA[26], BesA[22], MexA[27] and MtrC all contain N-terminal cysteine residues which undergo lipidation to covalently attach to fatty acids of the inner membrane.[28-30] In the cases of MacA[25] and EmrA,[21] the N-terminal amino acids comprise a TM α-helix, which anchors these proteins to the inner-membrane. Naïvely, anchoring the adaptor protein to the inner membrane should remove a degree of freedom, restricting its motion to the plane of the RND transporter and increasing the likelihood of the adaptor-

transporter interaction. In fact, membrane attachment has been proposed to augment the adaptor-transporter complex in both AcrA and MexA.[31-33] We recently decided to explore this hypothesis, using two constructs: the full-length MtrC (MtrC FL) and MtrC with the native N-terminal replaced by that of AcrA (AcrASP-MtrC) (Fig. 10b and c).

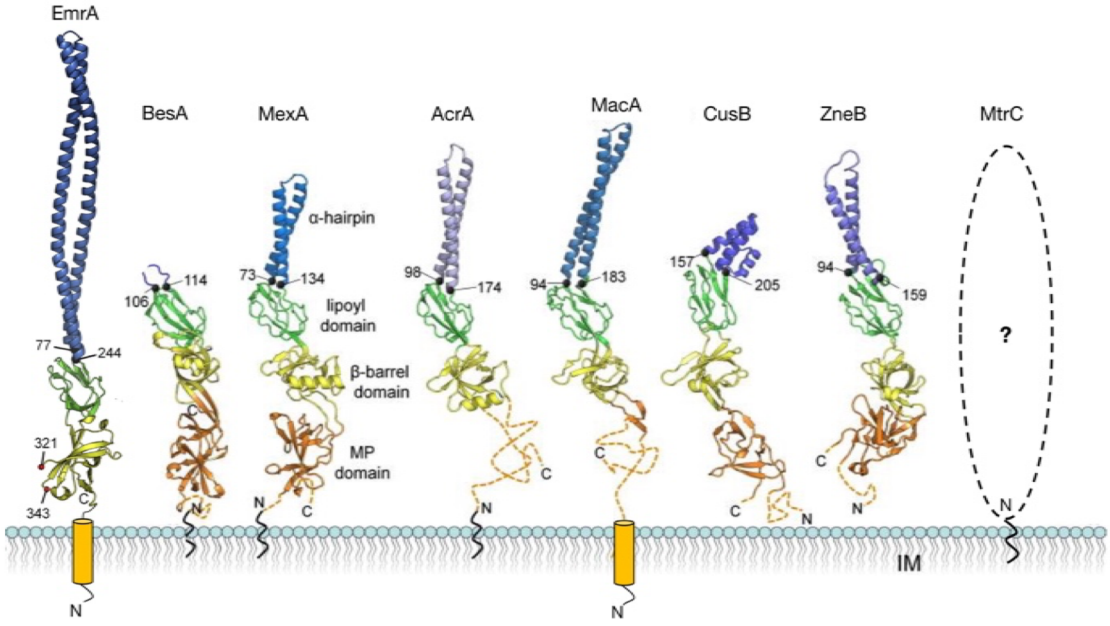

Figure 9. *Structural comparison of periplasmic adaptor proteins EmrA (Aquifex aeolicus), BesA (Borrelia burgdorferi), MexA (P. aeruginosa), AcrA (E. coli), MacA (E. coli macrolide efflux pump), and MtrC (N. gonorrhoeae) and the metal efflux pump adaptors CusB (E. coli) and ZneB (C. metallidurans). N-terminal domains of BesA, MexA, AcrA, and MtrC have an N-terminal lipoyl attachment site anchoring the adaptor in the IM, while EmrA and MacA are anchored by one TM helix. Dotted orange lines indicate unobserved MP domain terminal regions. Figure and caption adapted from Greene et al.*[22]

We were able to express each of these proteins in *E. coli* BL21ΔB(DE3), which harbors a genomic deletion of the *E. coli* efflux pump gene *acrB*. Unsurprisingly, these proteins were not detected in the soluble fraction of the cell lysate, but in the membrane fraction (data not shown). This result is distinct from that of MtrCΔNT, which was always recovered in the soluble fraction, supporting the role of the N-terminal residues as an anchor to the inner membrane.

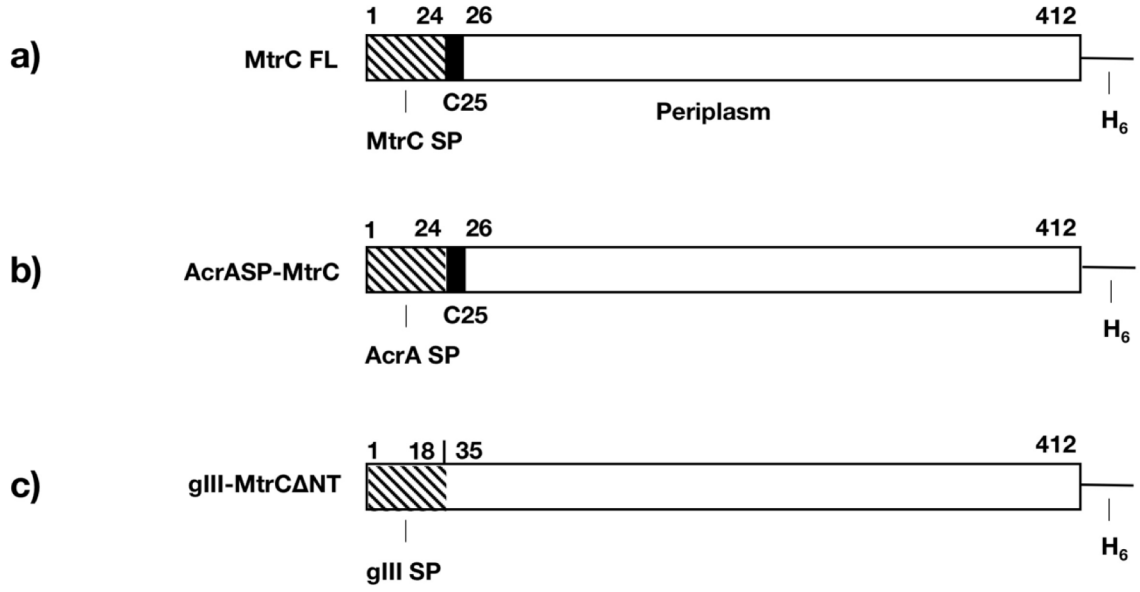

Figure 10. *Schematic depiction of the MtrC constructs used to assemble the tripartite MtrCDE pump. a) The full-length MtrC protein includes 24 N-terminal residues which target it to the periplasm of N. gonorrhoeae. Residue C25 is linked directly to fatty acids in the inner-membrane. b) The first 24 residues of the full-length MtrC construct were replaced by the corresponding residues of AcrA, which target it to the periplasm of E. coli. c) The first 34 residues of the full-length MtrC construct, including the lipidation site C25, were replaced by the 18-residue E. coli periplasmic signaling peptide gIII.*

Further, in large-scale purification experiments, we were able to recover well-folded, pure MtrC-FL protein from detergent solubilization of the *E. coli* cell membrane. We analyzed this purified protein sample using SEC to confirm that it was well-folded and homogeneous. Surprisingly, while MtrCΔNT eluted as a single peak at low molecular weight, MtrC FL predominantly formed a very high molecular weight peak, with a secondary peak roughly corresponding to the elution volume of MtrCΔNT (Fig. 11). These results indicate that, although a fraction of MtrC FL remains a monomer in solution, it tends toward a certain higher molecular weight species. As a comparison, we also analyzed purified MtrD protein in an identical fashion. Surprisingly, this protein eluted at almost an identical volume as MtrC FL, indicating that the molecular weight of the MtrD trimer is similar to the

molecular weight of the higher MtrC FL species (Fig. 11). For comparison, the theoretical molecular weight of the MtrD trimer is approximately 338.6 kDa, while hexameric MtrC FL is approximately 261.6 kDa. Our hypothesis remained that the hexameric MtrC FL would more tightly associate to the MtrD trimer. However, when the purified MtrC FL and MtrD samples were mixed in a ratio of 2:1, respectively, the elution volume of the predominant peak remained unchanged (Fig. 11). If these two proteins were indeed associating with each other, we were not able to detect the complex using this technique.

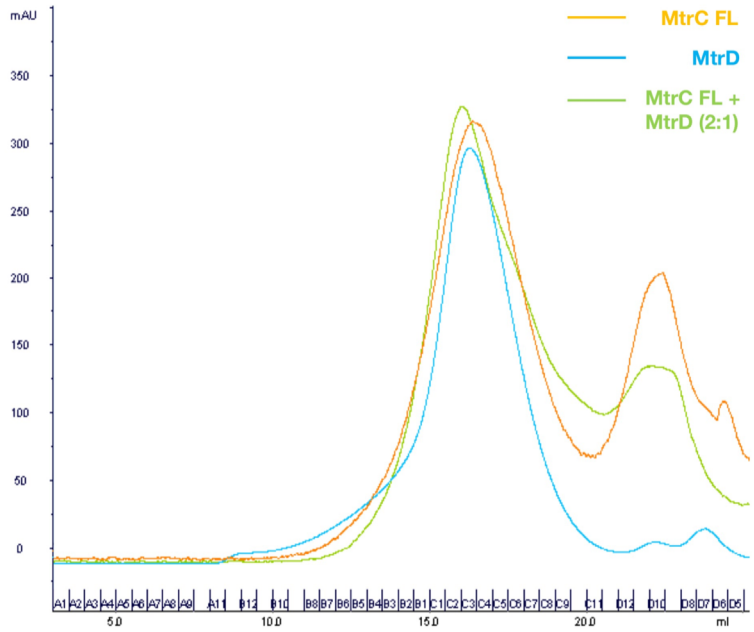

Figure 11. *Gel filtration experiment for the molecular size of MtrC FL and MtrD. Representative SEC analyses of MtrC FL (orange), MtrD (cyan), and a 2:1 (C:D) molar ratio mixture of MtrC FL and MtrD (green). The results shown are for the homologous CmeABC efflux system of C. jejuni, but similar results were obtained for MtrC and MtrD.*

We also analyzed the antimicrobial efflux activity of these proteins *in vivo*, using a minimum inhibitory concentration (MIC). In the MIC assay, serial dilutions of antimicrobials are incorporated into agar media to determine the lowest concentration of antimicrobial that will still permit bacterial growth. If the MtrCDE tripartite efflux system confers resistance to a particular toxin, then cells expressing these proteins should be able to tolerate a higher concentration compared to cells that do not contain the MtrCDE proteins.

We induced expression of the MtrCΔNT, MtrC FL, and AcrASP-MtrC proteins, along with MtrD and MtrE in each case, in *E. coli* BL21ΔB(DE3), to remove the background resistance of these cells conferred by the AcrB RND efflux protein. As the MtrCΔNT protein does not contain a SP sequence and, therefore, would not be targeted to the periplasm in this experiment, the protein was modified with an N-terminal gIII signal peptide, which is native to *E. coli* and should allow efficient secretion of MtrCΔNT to the periplasm (Fig. 10c). As expected, western blot analysis revealed that this gIII-MtrCΔNT protein was found in the soluble fraction, while MtrC FL was only found in the membrane fraction (Fig. 12). *E. coli* cells harboring these proteins, as well as a control sample that contained only the empty expression vector, were plated onto increasing concentrations of TX-100, a non-ionic detergent and known substrate of the Mtr system.[34,35] The results were striking. While *E. coli* cells harboring the empty vector were only able to grow to concentrations of 40 μg/mL TX-100, cells expressing gIII-MtrCΔNT tolerated concentrations exceeding 2560 μg/mL (Fig. 13). More surprising is that cells expressing either MtrC FL or AcrASP-MtrC

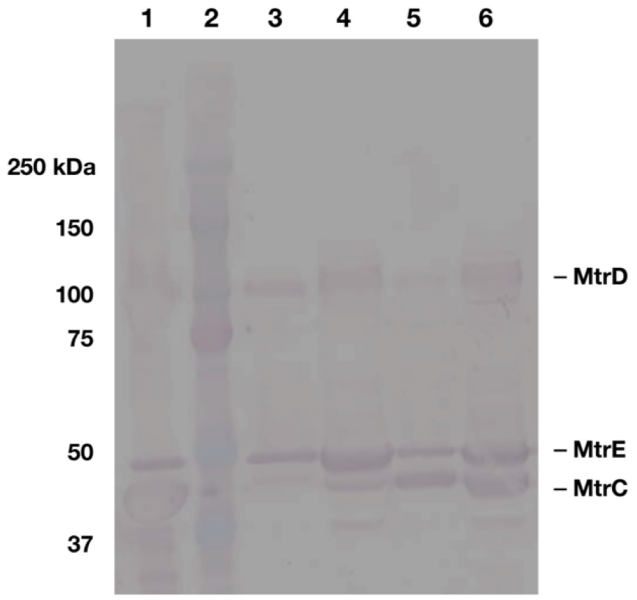

Figure 12. *Western blot analysis of MtrD, MtrE, and MtrC constructs co-expressed for MIC assays. Lane 1, MtrCDE positive control. Lane 2, protein molecular weight standards. Lane 3, MtrC FL-DE coexpression, soluble fraction. Lane 4, MtrC FL-DE coexpression, membrane fraction. Lane 5, gIII-MtrCΔNT-DE coexpression, soluble fraction. Lane 6, gIII-MtrCΔNT-DE coexpression, membrane fraction.*

were even less resistant than cells harboring the empty vector, growing only at the lowest concentration of TX-100 (20 μg/mL).

Contrary to the results of Yoneyama *et al*.,[30] which indicate that lipid modification of the membrane fusion protein is not essential to the operation of the pump, our results indicate that it may actually impair it. This cannot be the case, as the membrane anchoring domain of the adaptor protein is a well-conserved feature across all tripartite efflux systems. Instead, it is possible that the native signal sequence of *N. gonorrhoeae* MtrC is not recognized by the *E. coli* cells we used in this experiment. Thus, we repeated the MIC experiment with the native SP of MtrC replaced by that of AcrA. Again, we obtained the same results: cells expressing the gIII-MtrCΔNT protein outperformed AcrASP-MtrC with an MIC greater by two orders of magnitude. It is possible that the engineering of MtrC did interfere with its function. In this case, the important result is that MtrC is indispensable for the pump function.

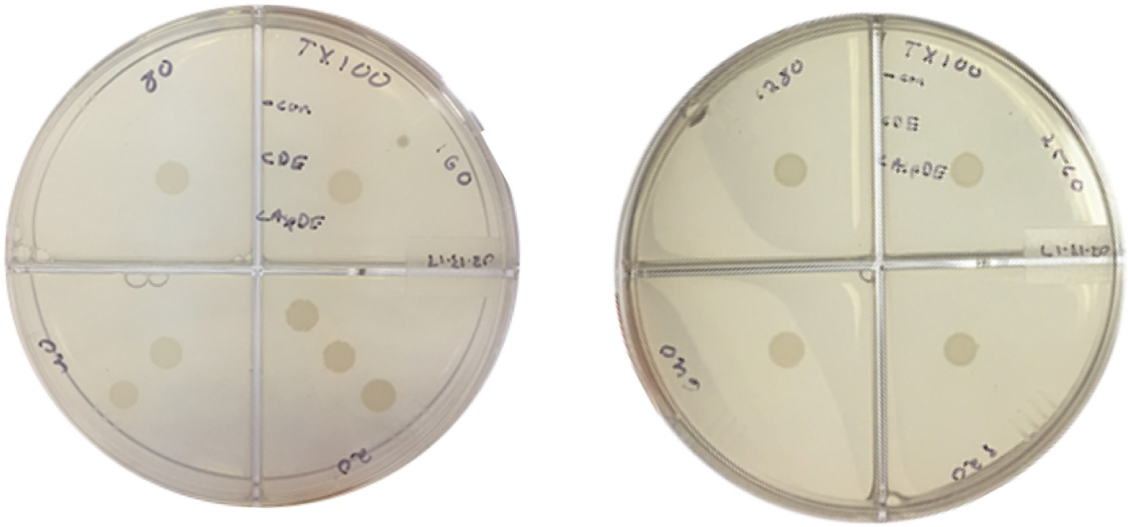

Figure 13. *MIC assay for the function of the MtrCDE efflux pump. Outer drop, empty vector. Middle drop, gIII-MtrCΔNT-MtrD-MtrE. Inner drop, AcrASP-MtrC-MtrD-MtrE. E. coli BL21ΔB(DE3) cells were used for all experiments. Identical results were observed for both AcrASP-MtrC-MtrD-MtrE and MtrC FL-MtrD-MtrE.*

Conclusion

We have successfully co-crystallized the bipartite MtrCD complex, which yielded a preliminary low resolution model. This model confirms the stoichiometry of 3:6 adaptor:transporter, which has been observed directly in the Acr and Mex systems. Using a combination of mutagenesis, chemical cross-linking, and protein engineering, we have developed first steps to increase the affinity of the MtrC, MtrD, and MtrE proteins for the tripartite complex. Using MIC assays, we have demonstrated the efficiency of the MtrCDE efflux system, which is capable of increasing the tolerance of *E. coli* cells for the antimicrobial TX-100 by two orders of magnitude. This result also demonstrates the importance of the membrane fusion protein to the function of the pump, as expression of a modified MtrC protein in *E. coli* even failed to restore the intrinsic bacterial resistance for TX-100. Thus, MtrC is a particularly attractive target for drugs to combat the growing prevalence of antibiotic resistant strains of *N. gonorrhoeae*. This task still necesitates a high-resolution structure of the assembled tripartite efflux complex. We are hopeful that the techniques presented in this chapter will contribute to this important work.

References

1. Higgins CF (2007). Multiple molecular mechanisms for multidrug resistance transporters. *Nature* 446:749-757.

2. Routh MD, Zalucki Y, Su CC, Long F, Zhang Q, Shafer WM, Yu EW (2011). Efflux pumps of the resistance-nodulation-division family: a perspective of their structure, function, and regulation in gram-negative bacteria. In *Advances in Enzymology and Related Areas of Molecular Biology*, Vol. 77, ed. EJ Toone, pp. 109-146. Hoboken, NJ: Wiley

3. Delmar JA, Su CC, Yu EW (2014). Bacterial multidrug efflux transporters. *Ann. Rev. Biophys.* 43:93-117

4. Shafer WM, Qu XD, Waring AJ, Lehrer RI (1998). Modulation of Neisseria gonorrhoeae susceptibility to vertebrate antibacterial peptides due to a member of the resistance/nodulation/division efflux pump family. *Proc. Natl. Acad. Sci. USA* 95:1829-1833

5. Hagman KE, Lucas CE, Balthazar JT, Snyder LA, Nilles M, Judd RC, Shafer WM (1997). The MtrD protein of Neisseria gonorrhoeae is a member of resistance/nodulation/division protein family constituting part of an efflux system. *Microbiology* 143:2117-2125

6. Hagman KE, Pan W, Spratt BG, Balthazar JT, Judd RC, Shafer WM (1995). Resistance of Neisseria gonorrhoeae to antimicrobial hydrophobic agents is modulated by the mtrRCDE efflux system. *Microbiology* 141:611-622

7. Delahay RM, Robertson BD, Balthazar JT, Ison CA (1997) Involvement of the gonococcal MtrE protein in the resistance of Neisseria gonorrhoeae to toxic hydrophobic agents. *Microbiology* 143:2127-2133

8. Lei HT, Chou TH, Su CC, Bolla JR, Kumar N, Radhakrishnan A, Long F, Delmar JA, Do SV, Rajashankar KR, Shafer WM, Yu EW (2014). Crystal structure of the open state of the Neisseria gonorrhoeae MtrE outer membrane channel. *PLoS One* 9(6):e97475

9. Su CC, Bolla JR, Kumar N, Radhakrishnan A, Long F, Delmar JA, Chou TH, Rajashankar KR, Shafer WM, Yu EW (2015). Structure and function of Neisseria gonorrhoeae MtrF illuminates a class of antimetabolite efflux pumps. *Cell Rep.* 11(1):67-70

10. Bolla JR, Su CC, Do SV, Radhakrishnan A, Kumar N, Long F, Chou TH, Delmar JA, Lei HT, Rajashankar KR, Shafer WM, Yu EW (2014). Crystal structure of the Neisseria gonorrhoeae MtrD inner membrane multidrug efflux pump. *PLoS One* 9(6):e97903

11. Delmar JA, Su CC, Yu EW (2015). Heavy metal transport by the CusCFBA efflux system. *Prot. Sci.* 24(11):1720-1736

12. Sennhauser G, Amstutz P, Briand C, Storchenegger O, Grütten MG (2007). Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS One* 5(1):e7

13. Janganan TK, Bavro VN, Zhang L, Matak-Vinkovic D, Barrera NP, Venien-Bryan C, Robinson CV, Borges-Walmsley MI, Walmsley AR (2011). Evidence for the assembly of a bacterial tripartite multidrug pump with a stoichiometry of 3:6:3. *J. Biol. Chem.* 286(30):26900-26912

14. Dijun D, Wang Z, James NR, Voss JE, Klimont E, Ohene-Agyei T, Venter H, Chiu W, Luisi BF (2014). Structure of the AcrAB-TolC multidrug efflux pump.

15. Jeong H, Kim JS, Song S, Shigematsu H, Yokoyama T, Hyun J, Ha NC (2016). Pseudoatomic structure of the tripartite multidrug efflux pump AcrAB-TolC reveals the intermeshing cogwheel-like interaction between AcrA and TolC. *Structure* 24(2):272-276

16. Daury L, Orange F, Taveau JC, Verchère A, Monlezun L, Gounou C, Marreddy RKR, Picard M, Broutin I, Pos KM, Lambert O (2016). Tripartite assembly of RND multidrug efflux pumps. *Nat. Commun.* 7:10731

17. Janganan TK, Bavro VN, Zhang L, Borges-Walmsley MI, Walmsley AR (2013). Tripartite efflux pumps: energy is required for dissociation, but not assembly or opening of the outer membrane channel of the pump. *Mol. Microbiol.* 88(3):590-602

18. Guan L, Nakae T (2001). Identification of essential charged residues in transmembrane segments of the multidrug transporter MexB of Pseudomonas aeruginosa. *J. Bacteriol.* 183(5):1734-1739

19. Su CC, Li M, Gu R, Takatsuka Y, McDermott G, Nikaido H, Yu EW (2006). Conformation of the AcrB multidrug efflux pump in mutants of the putative proton relay pathway. *J. Bacteriol.* 188(20):7290-7296

20. Seeger MA, von Ballmoos C, Verrey F, Pos KM (2009). Crucial role of Asp408 in the proton translocation pathway of multidrug transporter AcrB: evidence from site-directed mutagenesis and carbodiimide labeling. *Biochemistry* 48(25):5801-5812

21. Hinchliffe P, Greene NP, Paterson NG, Crow A, Hughes C, Koronakis V (2014). Structure of the periplasmic adaptor protein from a major facilitator superfamily (MFS) multidrug efflux pump. *FEBS Lett.* 588(17):3147-3153

22. Greene NP, Hinchliffe P, Crow A, Ababou A, Hughes C, Koronakis V (2013). Structure of an atypical periplasmic adaptor from a multidrug efflux pump of the spirochete Borrelia burgdorferi. *FEBS Lett.* 587(18):2984-2988.

23. De Angelis F, Lee JK, O'Connell JD III, Miercke LJW, Verschueren KH, Srinivasan V Bauvois C, Govaerts C, Robbins RA, Ruysschaert JM, Stroud RM, Vandenbussche G (2010). Metal-induced conformational changes in ZneB suggest an active role of membrane fusion proteins in efflux resistance systems. *Proc. Natl. Acad. Sci. USA* 107(24):11038-11043

24. Su CC, Yang F, Long F, Reyon D, Routh MD, Kuo DW, Mokhtari AK, Van Ornam JD, Rabe KL, Hoy JA, Lee YJ, Rajashankar KR, Yu EW (2009). Crystal structure of the membrane fusion protein CusB from Escherichia coli. *J. Mol. Biol.* 393(2):342-355

25. Yum S, Xu Y, Piao S, Sim SH, Kim HM, Jo WS, Kim KJ, Kweon HS, Jeong MH, Jeon H, Lee K, Ha NC (2009). Crystal structure of the periplasmic component of a tripartite macrolide-specific efflux pump. *J. Mol. Biol.* 387(5):1286-1297

26. Mikolosko J, Bobyk K, Zgurskaya HI, Ghosh P (2006). Conformational flexibility in the multidrug efflux system protein AcrA. *Structure*. 14:577-587

27. Akama H, Matsuura T, Kashiwagi S, Yoneyama H, Narita S, Tsukihara T, Nakagawa A, Nakae T (2004). Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in Pseudomonas aeruginosa. *J. Biol. Chem.* 279:25939-25942

28. Dinh T, Paulsen IT, Saeir MH Jr (1997). A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of gram-negative bacteria. *J. Bacteriol.* 176(13):3825-3831

29. Paulsen IT, Park JH, Choi PS, Saier MH Jr (1997) A family of Gram-negative bacterial outer membrane factors that function in the export of proteins, carbohydrates, drugs and heavy metals from Gram-negative bacteria. *FEMS Microbiol. Lett.* 156(1):1-8

30. Yoneyama H, Maseda H, Kamiguchi H, Nakae T (2000). Function of the membrane fusion protein, MexA, of the MexA, B-OprM efflux pump in Pseudomonas aeruginosa without an anchoring membrane. *J. Biol. Chem.* 275(7):4628-4634

31. Tikhonova EB, Yamada Y, Zgurskaya HI (2011). Sequential mechanism of assembly of multidrug efflux pump AcrAB-TolC. *Chem. Biol.* 18(4):454-463

32. Enguéne VYN, Verchère A, Phan G, Broutin I, Picard M (2015). Catch me if you can: a biotinylated proteoliposome affinity assay for the investigation of assembly of the MexA-MexB-OprM efflux pump from Pseudomonas aeruginosa. *Front. Microbiol.* 6:541

33. Ferrandez Y, Monlezun L, Phan G, Benabdelhak H, Benas P, Ulryck N, Falson P, Ducruix P, Picard M, Broutin I (2012). Stoichiometry of the MexA-OprM binding, as investigated by blue native gel electrophoresis. *Electrophoresis* 33:1282-1287

34. Rouquette C, Harmon JB, Shafer WM (1999). Induction of the mtrCDE-encoded efflux pump system of Neisseria gonorrhoeae requires MtrA, an AraC-like protein. *Mol. Microbiol.* 33(3):651-658

35. Folster JP, Shafer WM (2005). Regulation of mtrF expression in Neisseria gonorrhoeae and its role in high-level antimicrobial resistance. *J. Bacteriol.* 187(11):3713-3720

CHAPTER V

COCRYSTAL STRUCTURE OF THE INTERMEMBRANE SPACE REGION OF THE PLASTID DIVISION PROTEINS PARC6 AND PDV1

Introduction

Like their cyanobacterial ancestors, chloroplast division in plants and algae proceeds by binary fission. The fission process involves the coordinated assembly of four rings both inside and outside the cell at the midplastid: the inner PD ring, the inner FtsZ ring (Z-ring), the outer PD ring and the outer DRP5B ring.[1,2] The simultaneous constriction of the stromal Z-ring and the periplasmic DRP5B ring, along with the inner and outer PD rings, pinches the chloroplast cell in two. Separation into daughter cells requires a large network of coordinating proteins.

Within this network, a pair of paralogous inner envelope membrane proteins, PARC6 and ARC6, and a pair of paralogous outer envelope membrane proteins, PDV1 and PDV2, are likely responsible for bridging both the inner and outer membranes to connect the Z-ring and DRP5B ring (Fig. 1).[3,4] Recent data indicates that these proteins have similar topologies, with a single transmembrane region and short C-terminal region residing in the intermembrane space (IMS).[3,5-7] In particular, PARC6 possesses a long stromal region that interacts with the FtsZ2 protein component of the inner Z-ring.[5] The IMS region of PARC6 has been shown to interact with the IMS region of PDV1.[4,5] Similar interactions have been observed between ARC6, FtsZ2 and PDV2.[3,5,8,9] Thus, the PARC6-PDV1 and ARC6-PDV2 complexes are capable of spanning both the inner (IEM) and outer (OEM) chloroplast membranes to coordinate the FtsZ and DRP5B rings during chloroplast cell division. Although PARC6 and ARC6 share significant sequence identity (24%), data suggests that

PARC6, which inhibits FtsZ assembly, acts antagonistically to ARC6, which promotes FtsZ assembly.[1,2,4,5,7,8,10]

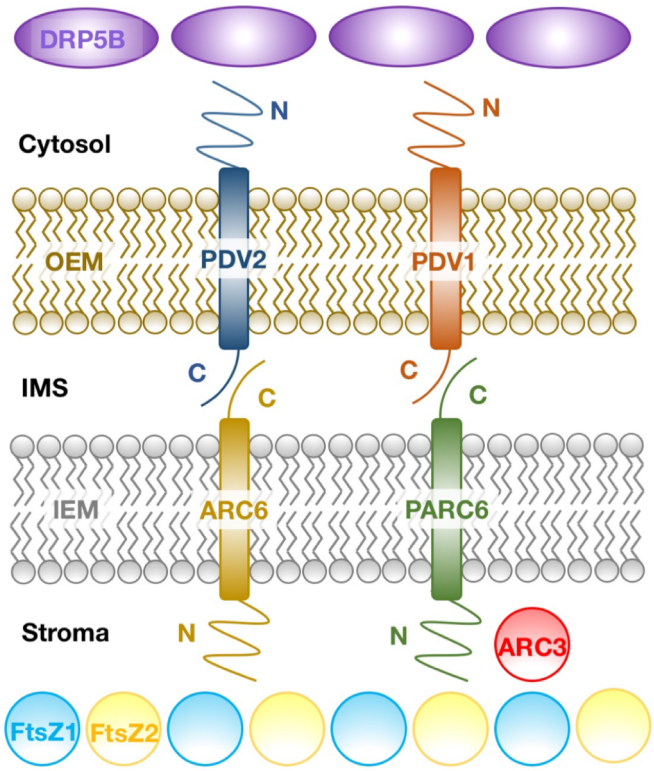

Figure 1. *Model of the chloroplast cell division machinery. The stromal proteins FtsZ1 and FtsZ2 and the cytosolic protein DRP5B form contractile rings. These two rings are coordinated by a pair of paralogous IEM proteins, ARC6 and PARC6, and a corresponding pair of paralogous OEM proteins, PDV1 and PDV2.*

While ARC6 is conserved throughout green-lineage chloroplasts, PARC6 is unique to vascular plants.[4,5] Like ARC6, PARC6 is a nuclear-encoded, IEM protein and a descendant of the cyanobacterial cell division protein Ftn2/ZipN.[7,11-13] Alignment of protein sequences suggests that *Arabidopsis thaliana* PARC6 and *Synechocystis* sp. Ftn2 share about 21% identity. The first 67 amino acids of *A. thaliana* PARC6 encode a cleavable chloroplast transit peptide.[7] The remainder of the protein can be divided into a larger N-terminal stromal-facing region (residues 77-573), transmembrane region (574-596) and smaller C-terminal IMS region (597-819).[5] A majority of the PARC6 IMS domain (residues 692-811) belongs to the highly-conserved Domain of unknown function (DUF4101) superfamily.

We report here the crystal structure of this highly conserved domain in the IMS portion of *A. thaliana* PARC6 (residues 684-819) at 2.52 Å resolution. This is the first structural characterization of the PARC6 protein. In addition, we have obtained the crystal structure of this PARC6 domain in complex with the IMS domain of *A. thaliana* PDV1, revealing the molecular details of the intermembrane space interaction during chloroplast cell division. Based on gel filtration experiments, PARC6 appears to form a dimer in solution, and that dimerization is dependent on interaction with PDV1. Using isothermal titration calorimetry, we find that the PARC6-PDV1 interaction is highly specific. It is likely that the DUF4101 domain and the C-terminal residues of the PDV proteins, especially the terminal glycine G272, are responsible for linking the stromal and cytosolic protein components of chloroplast cell division, bridging both the inner Z-ring and outer DRP5b-ring to coordinate chloroplast cell division.

Structure of PARC6-IMS$_{684\text{-}819}$

Initial efforts to crystallize the full-length IMS region of PARC6 (residues 596-819) using hanging-drop vapor diffusion were unsuccessful. However, long bar-shaped crystals were obtained when the protein was pretreated with trypsin. The crystal structure of residues 684-819 of PARC6 (referred to hereafter as PARC6-IMS$_{684\text{-}819}$) was determined to a resolution of 2.52 Å with excellent geometric qualities, using molecular replacement. The structure of the paralogous ARC6-IMS$_{667\text{-}799}$ (PDB ID: 5D9R)[11] was utilized as a search model. Presumably, trypsin cleaved a substantial portion of the PARC6 N-terminal, as we were not able to model the first 88 N-terminal amino acids. The final crystal structure was refined to R_{work} and R_{free} of 18.1% and 25.9%, respectively (Table 1, Figure 2).

Table 1. Data collection, phasing and structural refinement statistics of PARC6-IMS$_{684-819}$ and PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$

Data set	PARC6-IMS$_{684-819}$	PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$
Data collection		
Wavelength (Å)	0.979	0.979
Space group	$C222_1$	P1
Cell constants (Å)		
a	44.53	61.23
b	124.55	80.19
c	130.05	84.07
α, β, γ (°)	90,90,90	88.36,80.99,83.49
Resolution (Å)	2.52 (2.60-2.52)	3.37 (3.50-3.37)
Completeness (%)	99.5 (98.4)	97.1 (87.0)
Total reflections	1,365,108	2,395,188
Unique reflections	12,616	23,564
$CC_{1/2}$(high resolution shell)	0.658	0.536
Redundancy	8.3 (7.0)	3.1 (2.5)
R_{pim} (%)	4.9 (35.5)	19.4 (72.8)
$\langle I/\sigma(I)\rangle$	12.1 (1.8)	4.2 (1.1)
Refinement		
R_{work} (%)	18.1	25.0
R_{free} (%)	25.9	31.8
B-factors		
Overall(Å^2)	61.96	56.23
Chain A/B	66.08/65.59	53.79/48.87
Chain C/D/E/F/G/H		48.83/55.74/59.45/61.27/64.02/58.69
Rms deviations		
Bond (Å)	0.010	0.006
Angles (^{o})	1.02	0.90
Ramachandran analysis		
Most favored (%)	97.6	96.5
Allowed (%)	2.0	3.3
Disallowed (%)	0.3	0.2

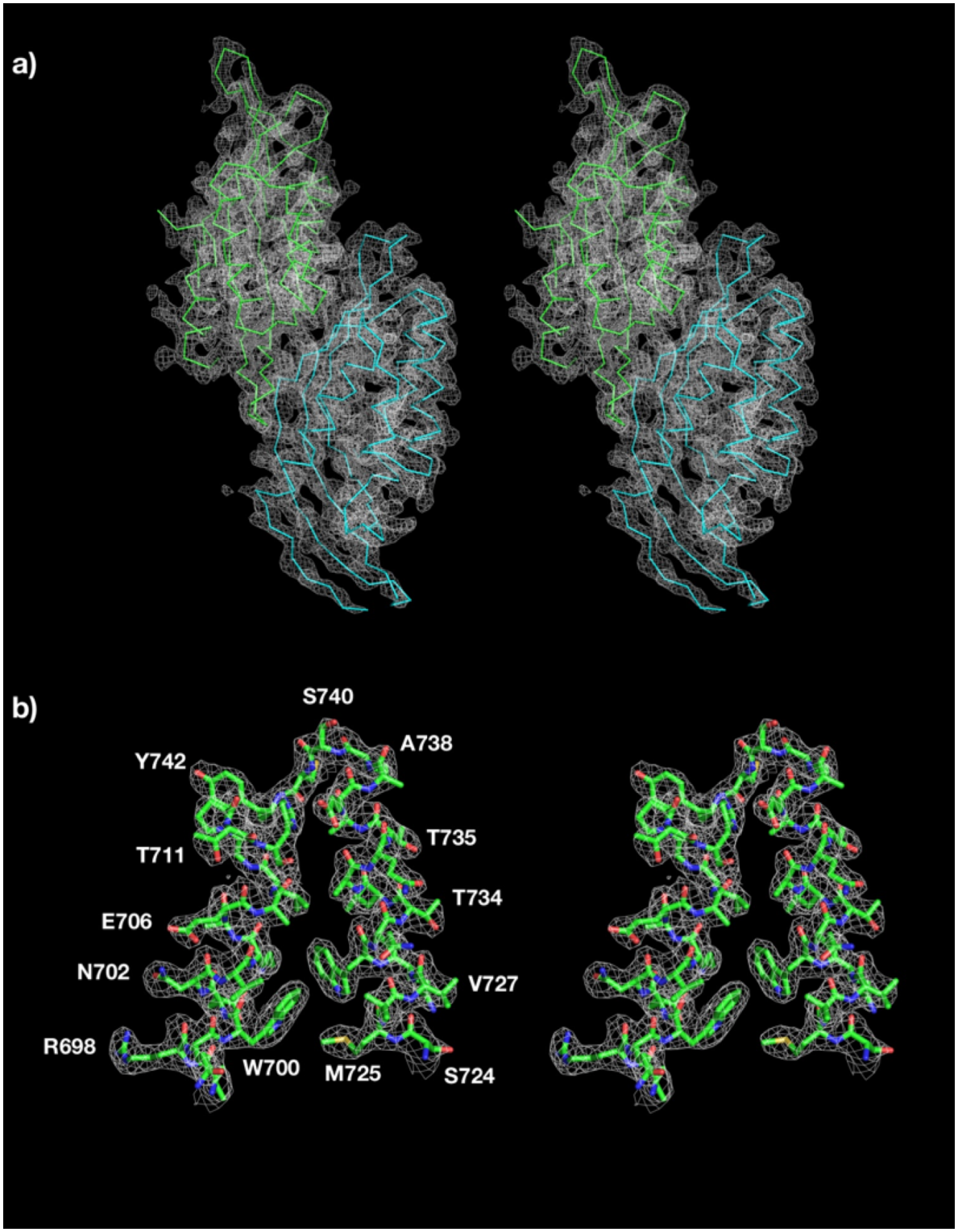

Figure 2. *Stereo view of the experimental electron density maps of PARC6-IMS$_{596-819}$ at a resolution of 2.52 Å. a) The electron density maps are contoured at 1.3 σ. The Cα traces of the two PARC6-IMS$_{596-819}$ molecules are colored green and cyan. b) Representative section of electron density in the vicinity of α1 and α3, within the α/β barrel. The solvent-flattened electron density (40-2.52 Å) is contoured at 1.3 σ and superimposed with the final refined model (green, carbon; red, oxygen; blue, nitrogen; yellow, sulfur).*

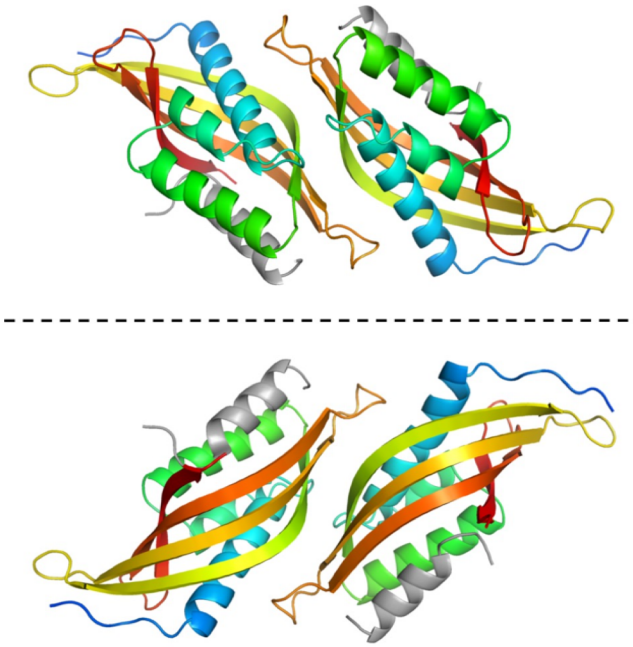

Figure 3. *Ribbon diagram of PARC6-IMS$_{684-819}$. Two molecules in the asymmetric unit are arranged as an anti-parallel dimer. Each molecule is colored using a rainbow gradient from the N-terminal (blue) to the C-terminal (red). The extra helix, corresponding to residues 641-658 of PARC6, is colored grey. The two views correspond to a 180° rotation about the axis shown (dotted line).*

Like ARC6-IMS$_{667-799}$, these three α-helices and four β-strands form an α/β barrel within each protomer. The bottom half of the barrel, adjacent to the IEM, is completely closed.

The interior wall of the α/β barrel is lined with several charged and aromatic residues, including W700, K704, W729, E775, and Y798, which are conserved among PARC6, ARC6 and

Two molecules of PARC6-IMS$_{684-819}$ were found in the asymmetric unit and appear to be arranged as an anti-parallel dimer (Figure 3). Otherwise, the structure of PARC6-IMS$_{684-819}$ is very similar to that of ARC6-IMS$_{636-801}$.[11] Each protomer is composed of three α-helices and four β-strands: α1 (690-708), α2 (714-720), α3 (724-739), β1 (742-760), β2 (763-780), β3 (788-803) and β4 (806-815 (Figure 5).

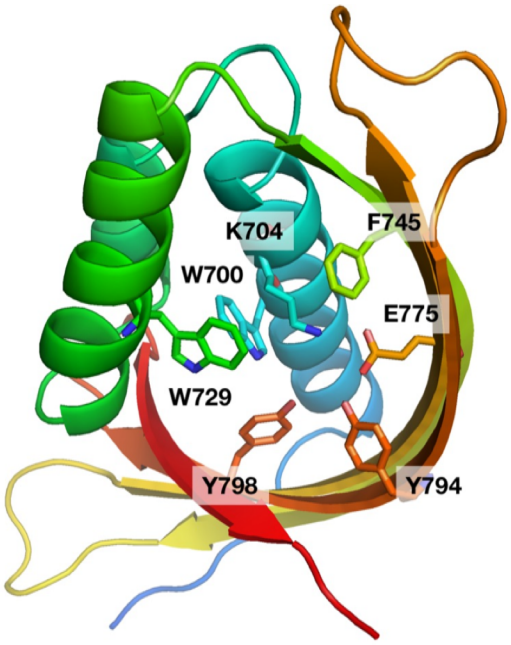

Figure 4. *Conserved amino acids within the PARC6-IMS$_{684-819}$ α/β barrel (rainbow, carbon; red, oxygen; blue, nitrogen). The secondary structural elements are colored using a rainbow gradient from the N-terminal (blue) to the C-terminal (red).*

Ftn2 proteins (Fig. 4). In addition, residue Y794, which is conserved between PARC6 and ARC6, but not Ftn2, is found within this α/β barrel. Interestingly, F745 is also conserved here, but only among PARC6 homologues. It is expected that these conserved residues are important for the function of this plastid division protein.

Surprisingly, an extra density, which forms an α-helical secondary structure, was discovered in the open face of the cavity created by the α/β barrel. We determined unambiguously that this density corresponds to residues 641-658 of PARC6, which was presumed to be removed from the IMS of PARC6 by trypsin digestion. The extra helix is secured atop the cavity by a disulfide bond between residues C657 and C741, which lies in the loop region between α3 and β1 (Fig. 5).

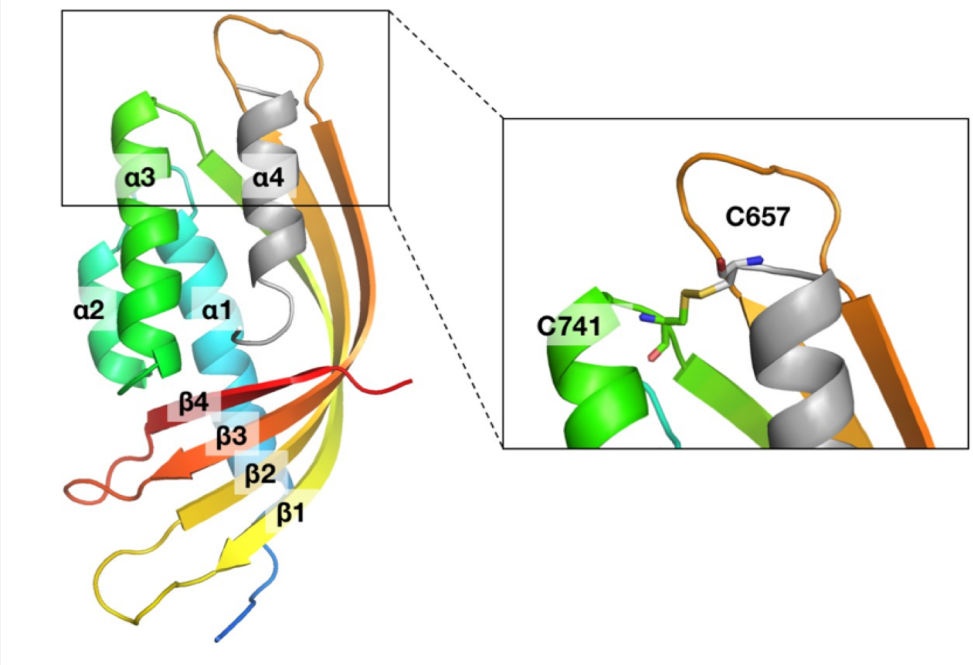

Figure 5. *Ribbon diagram of a single protomer of PARC6-IMS$_{684-819}$. Each molecule comprises three α-helices and four β-strands: α1 (residues 690-708), α2 (714-720), α3 (724-739), β1 (742-760), β2 (763-780), β3 (788-803) and β4 (806-815). The secondary structural elements are colored using a rainbow gradient from the N-terminal (blue) to the C-terminal (red). Helix α4 (642-657), colored grey, is secured atop the PARC6-IMS$_{684-819}$ α/β barrel by a disulfide bond between C657 and C741 (rainbow, carbon; red, oxygen; blue, nitrogen; yellow, sulfur).*

Co-crystal Structure of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$

We suspected that the cavity formed by the PARC6-IMS α/β barrel could create an ideal interaction site for the IMS domain of the PDV1 outer envelope membrane protein. To understand how PARC6 and PDV1 interact, we decided to co-crystallize the IMS domains of PARC6 and PDV1. We first expressed and purified the PARC6-IMS$_{684-819}$ protein that contains the C-terminal residues 684-819 of PARC6, elucidated by our previous crystal structure. We also produced the purified PDV1-IMS$_{225-272}$, which comprises the predicted IMS domain of PDV1 (Fig. 1). Initially, we mixed these two purified chloroplast proteins at a molar ratio of 1:1 for co-crystallization trials; unfortunately, these attempts were unsuccessful. We then made a construct that co-expresses PARC6-IMS$_{684-819}$ and PDV1-IMS$_{248-272}$. This construct contains a 6x(Gly-Ser) flexible linker that directly connects residue 819 of PARC6 to residue 248 of PDV1 to produce the PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ protein complex (Fig. 6). The PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ fusion protein was expressed, purified and subjected to crystallization using vapor diffusion.

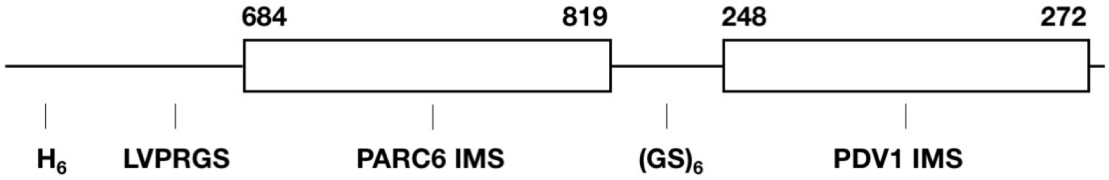

Figure 6. *Schematic depiction of the construct used for co-crystallization of PARC6-IMS and PDV1-IMS. The stable core of PARC6-IMS identified by crystallization (PARC6-IMS$_{684-819}$) was linked directly to the C-terminal 25 amino acids of PDV1 (PDV1-IMS$_{248-272}$) by a flexible 6xGly-Ser linker to create PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$. The PARC6-PDV1 chimera was tagged with 6xHis at the N-terminal and included a thrombin restriction site (LVPRGS) for removal of the tag.*

The structure of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ was determined to a final resolution of 3.37 Å. We utilized our structure of PARC6-IMS$_{684-819}$ as a search model for molecular replacement to obtain the crystal structure of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$.

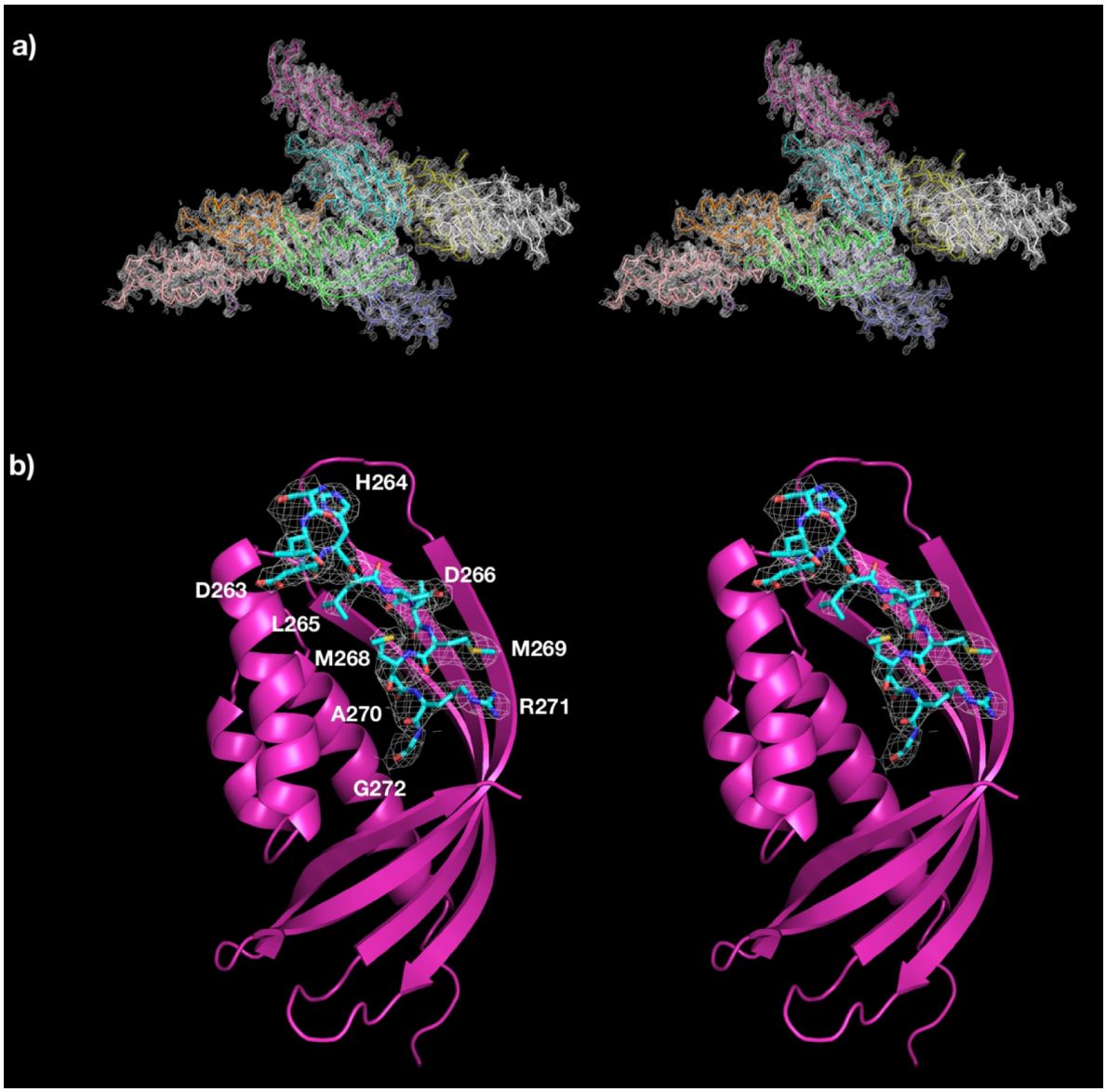

Figure 7. *Stereo view of the experimental electron density maps of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ at a resolution of 2.52 Å. a) The electron density maps are contoured at 1.3 σ. The Cα traces of the eight PARC6-IMS$_{596-819}$ molecules are colored green, cyan, magenta, yellow, pink, white, blue, and orange. b) Representative section of electron density in the vicinity of PDV1. The solvent-flattened electron density (40-3.37 Å) is contoured at 1.3 σ and superimposed with the final refined model (cyan, carbon; red, oxygen; blue, nitrogen; yellow, sulfur). The cartoon model of PARC6 is colored magenta.*

Eight molecules of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ are found in the asymmetric unit, with dimeric arrangements quite similar to that of PARC6-IMS$_{684-819}$ (Fig. 7a, Fig. 8). Superimposition of these eight molecules gives root mean square deviations (RMSDs) between 0.6 and 0.8 Å over 139-146 Cα atoms.

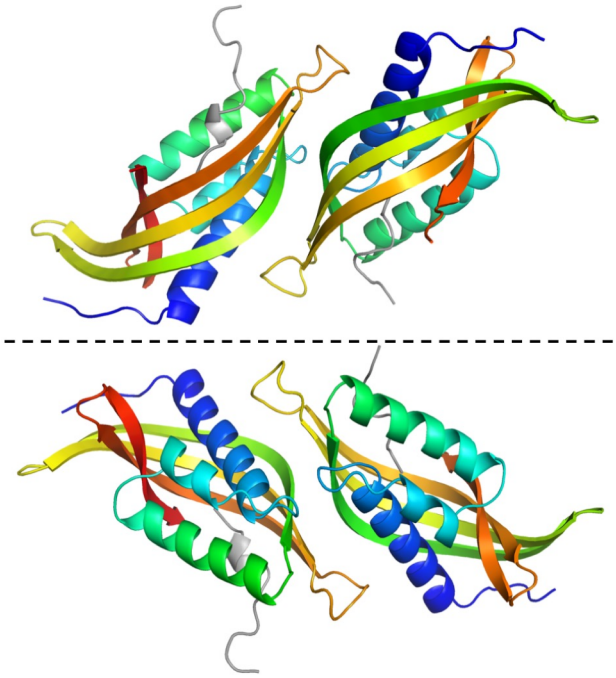

Figure 8. *Ribbon diagram of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$. Two molecules in the asymmetric unit are arranged as an anti-parallel dimer. Each molecule is colored using a rainbow gradient from the N-terminal (blue) to the C-terminal (red). The residues corresponding to PDV1-IMS are colored grey. The two views correspond to a 180° rotation about the axis shown (dotted line).*

Indeed, we found the C-terminal residues of PDV1-IMS stuck into the cavity created by the PARC6-IMS α/β barrel. At least twelve PDV1 amino acids, including residues 261-272, are visible in the crystal structure (Fig. 7b). These C-terminal residues of PDV1 mainly form a random coil (Fig. 9). Interestingly, the location of this PDV1 C-terminal random coil is found to overlap with that of the extra helix α4 (residues 641-658) of our previous PARC6 structure. Based on the PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ complex structure, the PDV1 C-terminal residue G272 is buried deep within the PARC6-IMS cavity. Previously, this residue was shown to be indispensable to the PARC6-PDV1 interaction.[5] This C-terminal glycine is also conserved between PDV1 and PDV2. It has been found that a single point mutation of these glycines to aspartic acid resulted in weak or no interaction with either PARC6 or

ARC6, respectively.[3]

Within hydrogen bonding distance of PDV1 G272 in our crystal structure is the conserved aromatic residue W700 of PARC6 (Fig. 9). The NE1 atom of this W700 is 2.6 Å away from the O atom of G272 of PDV1, on average, interacting to form a hydrogen bond.

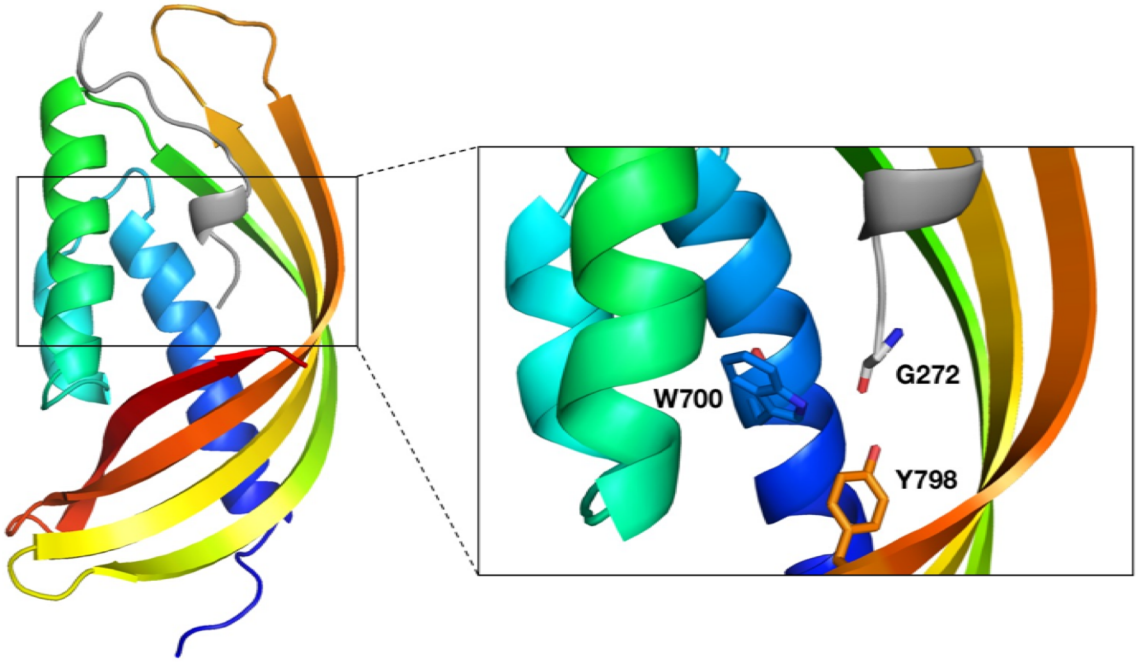

Figure 9. *Ribbon diagram of a single protomer of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$. Consistent with the structure of PARC6-IMS$_{684-819}$, each molecule comprises three α-helices and four β-strands. The secondary structural elements are colored using a rainbow gradient from the N-terminal (blue) to the C-terminal (red). The C-terminal glycine (G272) of PDV1-IMS is buried deep within the PARC6-IMS α/β barrel. Within hydrogen-bonding distance of G272 carboxyl group are the highly conserved residues W700 and Y798 of PARC6-IMS (rainbow, carbon; red, oxygen; blue, nitrogen).*

PARC6-PDV1 Interaction

To determine the strength of interaction between PARC6 and PDV1, we used isothermal titration calorimetry (ITC). The IMS region of PDV1, including residues 225-272 (PDV1-IMS$_{225-272}$), was purified alone and titrated into solution containing PARC6-IMS$_{684-}$

$_{819}$. This titration is characterized by a negative enthalpic contribution and yields a typical hyperbolic binding curve indicative of a specific interaction between these two chloroplast proteins (Fig. 10). The enthalpic (ΔH) and entropic (ΔS) parameters of binding of PDV1-IMS$_{225\text{-}272}$ to PARC6-IMS$_{684\text{-}819}$ are -9521 ± 60 cal·mol^{-1} and -0.87 cal·mol·deg^{-1}, giving rise to a dissociation constant, K_D, of 0.16 ± 0.01 μM.

According to our co-crystal structure, the PARC6 residue W700 was found in close proximity to the essential PDV1 C-terminal glycine.[5] We strongly suspect that this tryptophan residue might be involved in securing the binding of PDV1 to PARC6. Therefore, we decided to create single-point mutants of PARC6-IMS$_{684\text{-}819}$ in order to similarly study their interaction with the IMS domain of PDV1 by ITC. Unfortunately, all attempts to mutate this residue resulted in an unstable PARC6 protein and the binding data is unavailable.

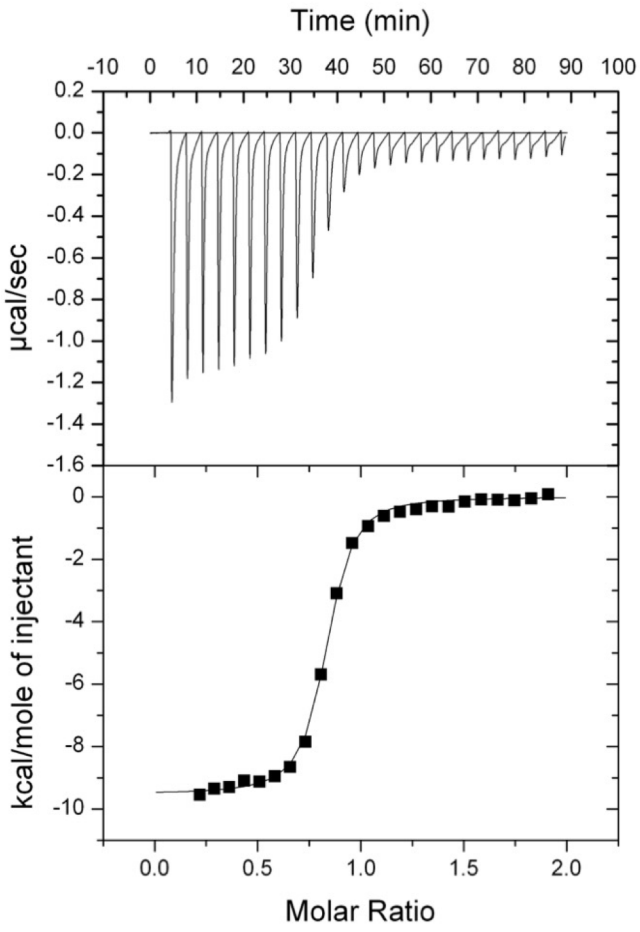

Figure 10. *Isothermal titration calorimetry for the binding of PARC6-IMS$_{684\text{-}819}$ and PDV1-IMS$_{225\text{-}272}$. Upper panel, each peak corresponds to the injection of 10 μL of 333 μM of PDV1-IMS$_{225\text{-}272}$ in buffer containing 20 mM Na-HEPES (pH 7.5), 250 mM NaCl, and 100 mM imidazole into the reaction containing 33 μM of PARC6-IMS$_{684\text{-}819}$ in the same buffer. Lower panel, cumulative heat of reaction is displayed as a function of the injection number. The solid line is the least square fit to the experimental data, giving a K_D of 0.16 ± 0.01 μM.*

PARC6 Dimerization

During chloroplast cell division, interactions between FtsZ proteins are necessary to form the Z ring. Preceding Z ring formation, these proteins have been shown not only to self-interact, forming homodimers, but proteins from each family of FtsZ (FtsZ1 and FtsZ2) can also interact to form heterodimers.[14] The interaction between FtsZ2 and the stromal domain of PARC6 is necessary for the proper division of chloroplast cells.[5]

To determine if PARC6 is similarly capable of dimerizing in solution, we employed analytical gel filtration experiments using the purified PARC-IMS$_{684\text{-}819}$, which comprises the conserved domain DUF4101 (Fig. 11a, 11b). The results suggest an average molecular weight of 15.8 ± 0.2 kDa (Fig. 12). This value is in agreement with the theoretical value of 16.8 kDa for one PARC-IMS$_{684\text{-}819}$ molecule, indicating that PARC-IMS$_{684\text{-}819}$ exists as a monomer in solution.

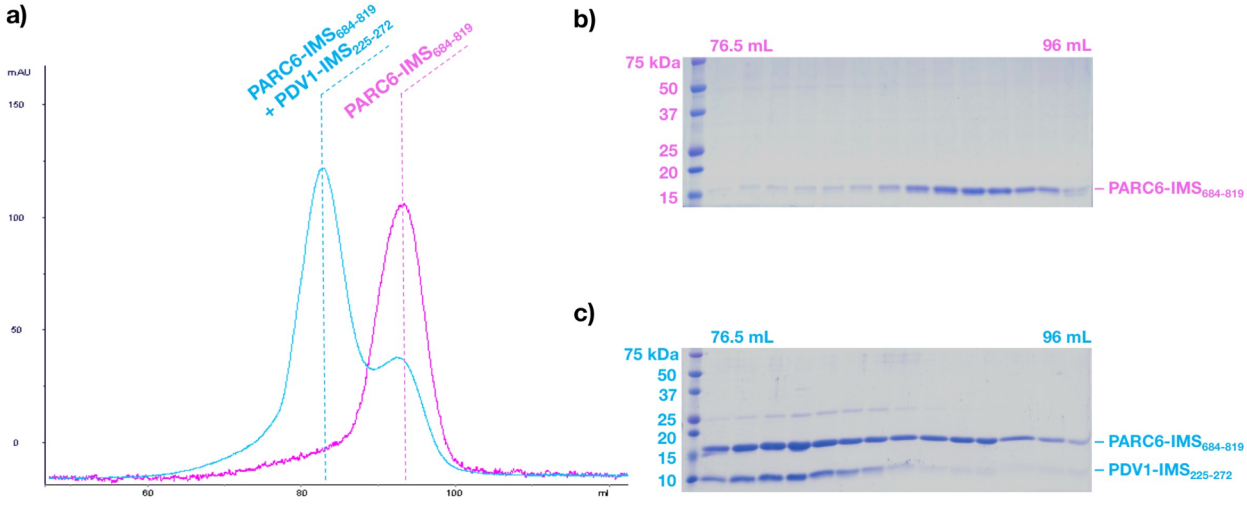

Figure 11. *Gel filtration experiment for the molecular size of PARC6-IMS$_{684\text{-}819}$ in the absence and presence of PDV1-IMS$_{225\text{-}272}$. a) Representative SEC analyses of PARC6-IMS$_{684\text{-}819}$ (magenta) and a 1:1 molar ratio mixture of PARC6-IMS$_{684\text{-}819}$ and PDV1-IMS$_{225\text{-}272}$ (cyan). b) and c) SDS-PAGE analysis of the indicated SEC peak fractions, respectively. The molecular mass of each marker protein (lane 1) is indicated on the left of each gel.*

To test if the interaction between PARC6 and PDV1 influences the oligomerization of PARC6, we also carried out gel filtration experiments with a mixture of the purified PARC6-IMS$_{684-819}$ and purified PDV1-IMS$_{225-272}$ in a 1:1 molar ratio (Fig. 11a, 11c). Suprisingly, the molecular weight of the PARC6 and PDV1 mixture was measured to be 37.0 ± 1.6 kDa (Fig. 12). The mass of the theoretically calculated dimer, including two molecules of PARC6-IMS$_{684-819}$ and two molecules of PDV1-IMS$_{225-272}$, is 44.8 kDa.

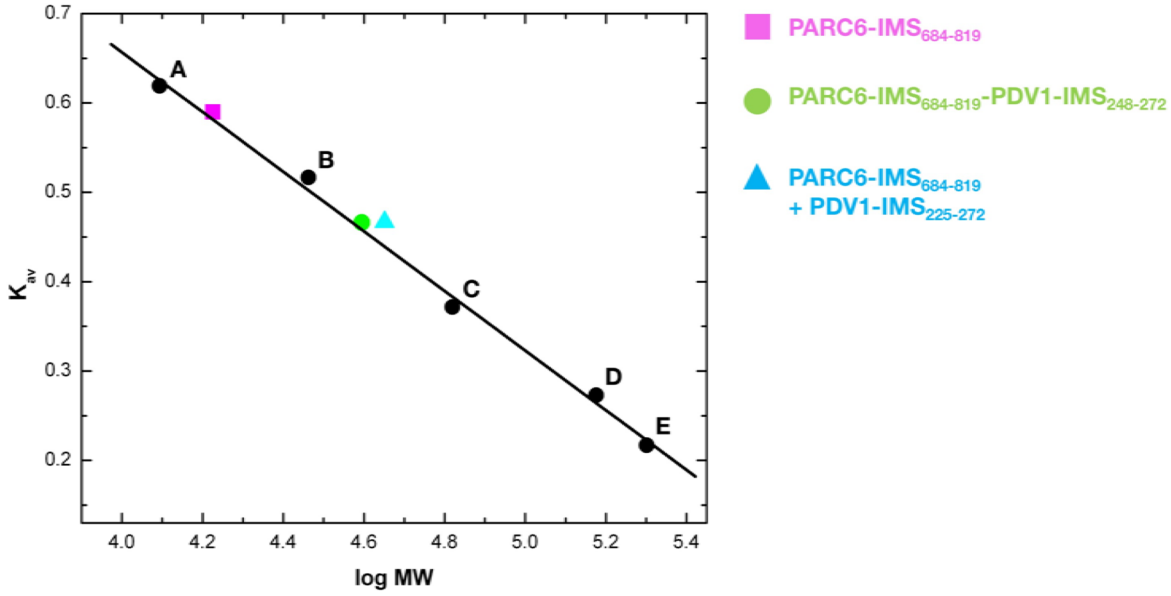

Figure 12. *Gel filtration experiments to determine molecular weight. The experiments demonstrate that while PARC6-IMS$_{684-819}$ (magenta square) is a monomer in solution, both PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ (green circle) and a 1:1 molar ratio mixture of PARC6-IMS$_{684-819}$ and PDV1-IMS$_{225-272}$ (cyan triangle) are dimers. The y-axis values were defined as: $K_{av} = (V_e - V_o)/(V_T - V_o)$, where V_T, V_e, and V_o are the total column volume, elution volume, and void volume of the column, respectively. Standards used were: A, cytochrome C (12,400 Da); B, carbonic anhydrase (29,000 Da); C, albumin bovine serum (66,000 Da); D, alcohol dehydrogenase (150,000 Da); and E, β-amylase (200,000 Da). The void volume was measured using blue dextran (MW 2,000,000).*

To test if the chimeric protein PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$, used for crystallization, is also capable of forming a dimer in solution, we carried out gel filtration experiments for this protein. The resulting elution profile was nearly identical to that of the 1:1 mixture of PARC6 and PDV1 (data not shown), with a measured mass of 37.2 ± 3.0 kDa.

This value is in excellent agreement with the theoretically calculated mass of 39.3 kDa for two molecules of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$, suggesting the PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ complex is dimeric in solution (Fig. 12).

The observed dimerization of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ is consistent with the dimeric arrangement observed in our crystal structure (Fig. 8). However, we observed an identical arrangement in our structure of the trypsin-digested PARC6-IMS$_{684-819}$ (Fig. 3). While there is no PDV1 protein present in this crystal, the extra helix α4, corresponding to residues 641-658 of PARC6, occupies a similar position in the α/β barrel binding pocket. Thus, we suspected that PARC6 dimerization may also be induced by interaction with α4.

To test this hypothesis, we used a semi-native SDS-PAGE protocol, in which SDS and reducing agents (in this case, DTT) are absent from the loading buffer. Thus, it is possible for native complexes to remain intact. When PARC6-IMS$_{596-819}$ was run in the presence of SDS and DTT, only bands corresponding to the theoretical weight of the PARC6-IMS monomer (26.5 kDa) could be observed (Fig. 13, *Lane 3*). However, when run with the semi-native protocol, upper bands appeared corresponding to the weight of the PARC6 dimer (Fig. 13, *Lane 4*). When digested with trypsin, a single band was observed, corresponding to the weight of residues 596-819 of PARC6 alone (Fig. 13, *Lane 6*). Interestingly, when this same protein was run with the semi-native protocol, we observed an upward shift in the lower band by several kDa (Fig. 13, *Lane 7*). As the helix α4 was shown in our crystal structure to be secured by a disulfide bond, we expect that the removal of DTT in the semi-native protocol allowed this helix to remain attached to the PARC6 protein. This shift was coupled to the appearance of an upper (dimeric) band in the gel. Consistent with

our crystal structures, it appears that the interaction between α4 and the PARC6-IMS α/β barrel may induce dimerization of this protein.

Taken together, our experiments suggest that the specific interaction between PARC6 and PDV1 may be essential for PARC6-IMS$_{684-819}$ dimerization. Interestingly, we were also able to observe dimerization which did not depend on PDV1, but on other residues within the IMS domain of PARC6-IMS. While the interaction between PARC6 and PDV1 has been shown to be necessary for chloroplast cell division,[4] our results suggest that the resulting PARC6 dimerization may also play a role in this process.

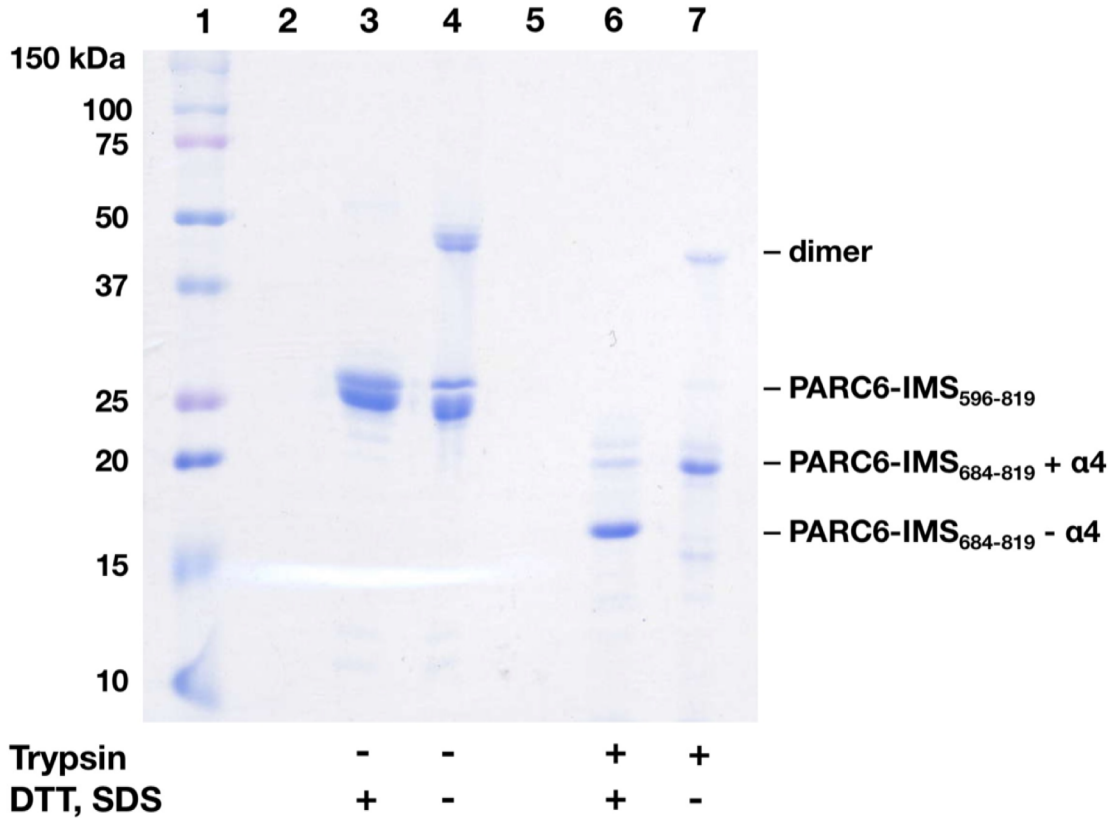

Figure 13. *Trypsin was added to purified PARC6IMS$_{596-819}$ at a molar ratio of 1:100, protease:PARC6IMS, with and without the addition of SDS and DTT, and the results were analyzed by SDS-PAGE. Lane 1, the molecular weight standard. Lane 2, empty. Lane 3, Purified PARC6IMS$_{596-819}$; no protease, with SDS and DTT. Lane 4, Purified PARC6IMS$_{596-819}$ no protease, without SDS and DTT. Lane 5, empty. Lane 6, Purified PARC6IMS$_{596-819}$ digested by trypsin, with SDS and DTT. Lane 7, Purified PARC6IMS$_{596-819}$ digested by trypsin, without SDS and DTT.*

Conclusion

Our crystal structures of PARC6-IMS encompass the highly conserved DUF4101 domain (residues 692-811), which spans most of the IMS domain of PARC6 and its paralog ARC6 (kumar 2016). Sequence alignment shows that the cyanobacterial cell division protein Ftn2 bears identical or similar residues at most of these positions, suggesting the importance of DUF4101 in the cell division processes of both cyanobacteria and its evolutionary descendant chloroplasts.[7,11] Interestingly, PARC6 and ARC6 share only 12% protein sequence identity over this region. The IMS regions of PARC6 and ARC6 interact with PDV1 and PDV2, respectively, and chloroplast cell division has been shown to rely on this interaction.[3-5]

Previously, it was not clear whether DUF4101 was necessary for the intermembrane space interaction during chloroplast cell division. Our experiments indicate that the DUF4101 domain of PARC6 and the C-terminus of PDV1, specifically the C-terminal glycine G272, are capable of mediating the interaction between these two proteins. Based on our crystal structures, the conserved aromatic residue PARC6 W700 is suspected to contact PDV1 G272 via a hydrogen bond, securing the binding between these two chloroplast division proteins. The fact that DUF4101 is also conserved in cyanobacterial Ftn2 proteins suggests that Ftn2 may function in a manner similar to the plant-specific PARC6 and ARC6 cell division machines.

Oligomerization in several chloroplast division-associated proteins has been shown to be necessary for their function. Specifically, FtsZ proteins can form homodimers as well as heterodimers among two protein families FtsZ1 and FtsZ2. It is possible that these interactions precede Z-ring formation during chloroplast division.[14] Recently, FtsZ2 proteins

have been found to interact directly with the stromal domain of PARC6.[5] According to our experimental data, the DUF4101 domain of PARC6 is a monomer in solution. However, when mixed or co-expressed with the C-terminal residues of PDV1, the dominant species is a dimer. Whether the PARC6 dimer we observe is important for the function of this protein in chloroplast cell division awaits further analysis. It is expected that our crystal structures, the first depicting interactions in the intramembrane space during chloroplast division, will facilitate further functional studies of these proteins and their roles in chloroplast and cyanobacterial cell division.

Methods

Expression and purification of PARC6-IMS$_{596-819}$

The plasmid pET15bΩ*PARC6-IMS*$_{596-819}$ bearing the PARC6-IMS region (amino acids 596-819) with a 6xHis tag at the N-terminus was overproduced in *E. coli* BL21(DE3) cells. Cells were grown in 6 L of Luria Broth (LB) medium with 100 μg/mL ampicillin at 37°C. When OD_{600} reached 0.4, the culture was treated with 0.2 mM isopropyl-β-d-thiogalactopyranoside (IPTG) to induce PARC6-IMS expression, and cells were harvested within 3 h. The cell pellet was suspended in 100 mL ice-cold buffer containing 20 mM Na-HEPES (pH 7.5) and 250 mM NaCl. The cells were then lysed with a French pressure cell. Cell debris was removed by centrifugation for 45 min at 4°C and 100,000 g. The crude lysate was filtered through a 0.2 μm membrane and was loaded onto a 5 mL Hi-Trap Ni^{2+}-chelating column (GE Healthcare Biosciences, Pittsburgh, PA) pre-equilibrated with 20 mM Na-HEPES (pH 7.5) and 250 mM NaCl. To remove unbound proteins and impurities, the column was first washed with six column volumes of buffer containing 50 mM imidazole,

250 mM NaCl and 20 mM Na-HEPES (pH 7.5). The PARC6-IMS protein was then eluted with four column volumes of buffer containing 300 mM imidazole, 250 mM NaCl, and 20 mM Na-HEPES (pH 7.5). The purity of the protein was judged using 15% SDS-PAGE stained with Coomassie Brilliant Blue. The purified protein was extensively dialyzed against buffer containing 20 mM Na-HEPES (pH 7.5), 250 mM NaCl and 100 mM imidazole and concentrated to 20 mg/mL. The procedures for purifying PARC6-IMS$_{684-819}$ were identical to those of PARC6-IMS$_{596-819}$.

Expression and purification of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$

The procedures for purifying PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ were identical to those of PARC6-IMS$_{596-819}$. Prior to crystallization of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272,}$ the N-terminal 6xHis tag was removed by thrombin digestion. 1 U of thrombin (Roche) was added to 10 mg of purified 6xHis PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ protein and incubated for 12 h at room temperature. The resulting tag-less PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ was further purified using size-exclusion chromatography in buffer containing 20 mM Na-HEPES (pH 7.5), 250 mM NaCl and 100 mM imidazole and concentrated to 20 mg/mL.

Expression and purification of PDV1-IMS$_{226-272}$

The plasmid pGEX4T1Ω *PDV1-IMS$_{226-272}$* bearing the PDV1-IMS region (amino acids 226-272) with a GST tag at the N-terminus was overproduced in *E. coli* BL21(DE3) cells. Cells were grown in 6 L of Luria Broth (LB) medium with 100 μg/mL ampicillin at 37°C. When OD_{600} reached 0.4, the culture was treated with 0.2 mM isopropyl-β-d-thiogalactopyranoside (IPTG) to induce PDV1-IMS expression, and cells were harvested

within 3 h. The cell pellet was suspended in 100 mL ice-cold buffer containing 20 mM Na-HEPES (pH 7.5) and 250 mM NaCl. The cells were then lysed with a French pressure cell. Cell debris was removed by centrifugation for 45 min at 4°C and 100,000 g. The crude lysate was filtered through a 0.2 μm membrane and was loaded onto a 5 mL Glutathione Sepharose High Performance column (GE Healthcare Biosciences, Pittsburgh, PA). To remove unbound proteins and impurities, the column was first washed with six column volumes of buffer containing 100 mM imidazole, 250 mM NaCl and 20 mM Na-HEPES (pH 7.5). The PDV1-IMS protein was then cleaved from the GST-bound beads by thrombin digestion. 1 U of thrombin (Roche) was added to the column per 10 mg of protein and incubated for 12 h at room temperature. The PDV1-IMS protein was then eluted with two column volumes of buffer containing 100 mM imidazole, 250 mM NaCl and 20 mM Na-HEPES (pH 7.5). The purity of the protein was judged using 18% SDS-PAGE stained with Coomassie Brilliant Blue.

Crystallization of PARC6-IMS$_{596-819}$

All crystals were grown using hanging-drop vapor diffusion. Crystals could only be obtained after removing the 6xHis tag and the first 88 N-terminal amino acids from PARC6-IMS$_{596-819}$ using trypsin. A 72 μg/mL final concentration of trypsin was added to 20 mg/mL purified PARC6-IMS$_{596-819}$. The protein mixture was incubated at 4°C for 30 min before crystallization, producing PARC6-IMS$_{684-819}$. For crystallization, 0.3 μL of the protein mixture was mixed with 0.3 μL of reservoir solution containing 18% polyethylene glycol (PEG) 4000, 0.1 M Magnesium acetate, 0.1 M Na-MES (pH 6.5), 0.05 M Potassium nitrate

and 5% glycerol. The resultant mixture was equilibrated against 500 μL of the reservoir solution. Bar-shaped crystals grew to full size within the drop within one week.

Crystallization of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$

For crystallization of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$, 0.3 μL of protein solution was mixed with 0.3 μL of reservoir solution containing 12% polyethylene glycol (PEG) 8000, 160 mM ammonium sulfate and 0.8% glycerol. The resultant mixture was equilibrated against 500 μL of the reservoir solution. Crystals grew to full size within the drop within one week.

Data collection, structural determination, and refinement of PARC6-IMS$_{684-819}$

All diffraction data were collected at 100K at beamline 24ID-C located at the Advanced Photon Source, using an ADSC Quantum 315 CCD-based detector. Diffraction data were processed using DENZO and scaled using SCALEPACK.[15] The crystals of PARC6-IMS$_{684-819}$ belong to space group C222$_1$ (Table 1). Based on the molecular weight (15.8 kDa), two monomers per asymmetric unit are expected with a solvent content of 56.8%. The phases of PARC6-IMS$_{684-819}$ were obtained using molecular replacement. The structure of the paralogous ARC6-IMS (residues 667-799) (PDB code 5D9R) was utilized as a template.[11] The resulting phases were then subjected to density modification and NCS averaging using the program RESOLVE.[16] The remaining part of the model was manually constructed using the program Coot.[17] The model was then refined using TLS refinement techniques adopting a single TLS body as implemented in PHENIX,[18] leaving 5% of reflections in Free-R set. Iterations of refinement using PHENIX[18] and model building in

Coot[17] led to the current structure, which consists of 134 residues with excellent geometrical characteristics (Table 1).

Data collection, structural determination, and refinement of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$

Diffraction data for crystals of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ were obtained and processed similar to that of PARC6-IMS$_{684-819}$. The crystals of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ belong to space group P1 (Table 1). Based on the molecular weight (19.5 kDa), eight monomers per asymmetric unit are expected with a solvent content of 55.15%. The phases of PARC6-IMS$_{684-819}$-PDV1-IMS$_{248-272}$ were obtained using molecular replacement. The structure of PARC6-IMS$_{684-819}$ was utilized as a template. The resulting phases were then subjected to density modification and NCS averaging using the program RESOLVE.[16] The remaining part of the model was manually constructed using the program Coot.[17] The model was then refined using TLS refinement techniques adopting a single TLS body as implemented in PHENIX,[18] leaving 5% of reflections in Free-R set. Iterations of refinement using PHENIX[18] and model building in Coot[17] led to the current structure, which consists of 143 residues with excellent geometrical characteristics (Table 1).

Isothermal titration calorimetry

All measurements were performed at 25°C using a VP-Microcalorimeter (MicroCal, Northampton, MA). Before titration, the protein concentrations were determined using the Bradford assay. To measure the interaction between PARC6-IMS and PDV1-IMS, the titrant solution contained the IMS domain of PDV1 (PDV1-IMS$_{226-272}$) at a concentration of 333

μM. The sample cell solution contained the IMS domain of PARC6-IMS (PARC6-IMS$_{684-819}$) at a concentration of 33 μM. The buffer for both the titrant and sample cell solutions contained 20 mM Na-HEPES (pH 7.5), 250 mM NaCl and 100 mM imidazole. Thirty 10 μL injections of the titrant solution were used for data collection. Injections occurred at intervals of 200 s and the duration of each injection was 10 s. Heat transfer ($\mu cal \cdot s^{-1}$) was measured as a function of elapsed time (s). The mean enthalpies measured from injection of the titrant in the buffer were subtracted from raw titration data before data analysis with ORIGIN software (MicroCal, Northampton, MA). Titration curves were fitted by a nonlinear least squares method to a function for the binding of a ligand to a macromolecule. Nonlinear regression fitting to the binding isotherm provided the equilibrium binding constant ($K_A = 1/K_D$) and enthalpy of binding (ΔH). Based on the values of K_A and ΔH, the change in free energy (ΔG) and entropy (ΔS) were calculated with the equation, $\Delta G = - RT \ln(K_A) = \Delta H - T\Delta S$, where T is 298 K and R is 1.9872 $cal \cdot K^{-1} \cdot mol^{-1}$. Calorimetry trials were also carried out in the absence of PDV1-IMS in the same experimental conditions. These heats of dilution were subtracted to produce the final data.

We also used isothermal titration calorimetry to determine the binding affinity of PARC6-IMS$_{684-819}$ W700A to PDV1-IMS$_{226-272}$. Experimental conditions were identical to those of the wild type proteins.

Gel filtration

A protein liquid chromatography Superdex 200 16/60 column (GE Healthcare Biosciences, Pittsburgh, PA) with a mobile phase containing 20 mM Na-HEPES (pH 7.5), 250 mM NaCl and 100 mM imidazole was used in the gel filtration experiments. Blue

dextran (Sigma-Aldrich, St. Louis, MO) was used to determine the column void volume and proteins for use as gel filtration molecular weight standards were cytochrome C (12,400 Da), carbonic anhydrase (29,000 Da), albumin bovine serum (66,000 Da), alcohol dehydrogenase (150,000 Da) and β-Amylase (200,000 Da). All standards were purchased from Sigma-Aldrich (St. Louis, MO). The molecular weights of the experimental samples were determined following the protocols supplied by the manufacturers.

References

1. Yang, Y., Glynn, J. M., Olson, B. J. S. C., Schmitz, A. J., Osteryoung, K. W. (2008) Plastid division: across time and space. *Curr. Opin. Plant Biol.* **11**, 577-584

2. Osteryoung, K. W., and Pyke, K. A. (2014) Division and dynamic morphology of plastids. *Annu. Rev. Plant Biol.* **65**, 443-472

3. Glynn, J. M., Froehlich, J. E., and Osteryoung, K. W. (2008) *Arabidopsis* ARC6 coordinates the division machineries of the inner and outer chloroplast membranes through interaction with PDV2 in the intermembrane space. *Plant Cell.* **20**, 2460-2470

4. Glynn, J. M., Yang, Y., Vitha, S., Schmitz, A. J., Hemmes, M., Miyagishima, S. Y., Osteryoung, K. W. (2009) PARC6, a novel chloroplast division factor, influences FtsZ assembly and is required for recruitment of PDV1 during chloroplast division in *Arabidopsis*. *Plant J.* **59**, 700-711

5. Zhang, M., Chen, C., Froehlich, J. E., TerBush, A. D., Osteryoung, K. W. (2016) Roles of *Arabidopsis* PARC6 in coordination of the chloroplast division complex and negative regulation of FtsZ assembly. *Plant Physiol.* **170**, 250-262

6. Miyagishima, S.–Y., Forehlich, J. E., Osteryoung, K. W. (2006) PDV1 and PDV2 mediate recruitment of the dynamin-related protein ARC5 to the plastid division site. *Plant Cell.* **18**, 2517-2530

7. Vitha, S., Froehlich, J. E., Koksharova, O., Pyke, K. A., van Erp, H., Osteryoung, K. W. (2003) ARC6 is a J-domain plastid division protein and evolutionary descendant of the cyanobacterial cell division protein Ftn2. *Plant Cell* **15**, 1918-1933

8. Zhang, M., Hu, Y., Jia, J., Li, D., Zhang, R., Gao, H., He, Y. (2009) CDP1, a novel component of chloroplast division site positioning system in *Arabidopsis*. *Cell Res.* **19**, 877-886

9. Maple, J., Aldridge, C., Møller, S. G. (2005) Plastid division is mediated by combinatorial assembly of plastid division proteins. *Plant J.* **43**, 811-823

10. Otteson, E., Zhong, R., Lamppa, G. K. (2010) Identification of a chloroplast division mutant coding for ARC6H, an ARC6 homolog that plays a nonredundant role. *Plant Sci.* **178**, 114-122

11. Kumar, N., Radhakrishnan, A., Su, C.-C., Osteryoung, K. W., Yu, E. W. (2016) Crystal structure of a conserved domain in the intermembrane space region of the plastid division protein ARC6. *Prot. Sci.* **25**, 523-529

12. Pyke, K. A., Rutherford, S. M., Robertson, E. J., Leech, R. M. (1994) *arc6*, a fertile *Arabidopsis* mutant with only two mesophyll cell chloroplasts. *Plant Physiol.* **106**, 1169-1177

13. Mazouni, K., Domain, F., Cassier-Chauvat, C., Chauvat, F. (2004) Molecular analysis of the key cytokinetic components of cyanobacteria: FtsZ, ZipN and MinCDE. *Mol. Microbiol.* **52**, 1145-1158

14. Aldridge, C., Maple, J., Møller, S. G. (2005) The molecular biology of plastid division in higher plants. *J. Exp. Bot.* **56**, 1061-1077

15. Otwinowski, Z., Minor, M. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. **276**, 307–326

16. Terwilliger, T. C. (2001) Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallogr.* **D57**, 1755-1762

17. Emsley, P., Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr.* **D60**, 2126

18. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCroy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr.* **58**, 1948-1954

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTIONS

A high-resolution protein crystal structure can allow us to picture how medicine interacts with its target; visualize conformational changes that determine what a protein does and how it does it; and rationally design better drugs or enzymes with enhanced activity. Nearly every pharmaceutical company harnesses this technique for drug discovery. Due to recent advances in robotic technology, the 'brute force' method of protein crystallization is increasingly viable. However, the biggest obstacle in membrane protein crystallography is not obtaining the crystals, but optimizing them for high-resolution X-ray diffraction. Therefore, it is necessary to have intimate knowledge of a particular protein's interaction with the detergent, precipitant, salt, and buffer present in a vapor diffusion crystallization experiment. The techniques outlined in Chapter I have allowed us to determine the high-resolution crystal structures of the outer membrane multidrug efflux channels of *Neisseria gonorrhoeae* MtrC and *Campylobacter jejuni* CmeC, the outer membrane metal ion efflux channel *Escherichia coli* CusC, the inner membrane RND multidrug efflux pumps *N. gonorrhoeae* MtrD and *E. coli* CusA, the first adaptor-transporter complex *E. coli* CusBA, and more.

Notably, we have obtained crystal structures of each of the proteins in the CusCBA tripartite efflux channel, allowing us to discern their mechanism of copper and silver efflux. This model is presented in Chapter II. Based on the crystal structure of CusA, it is possible that the methionine ladder created by M410 and M501, M403 and M486, and M391-M1009 could shuttle Cu^{+} and Ag^{+} ions from the cytoplasm to the primary three-methionine binding

site created by M573, M623, and M672. The periplasmic cleft of CusA containing these residues presumably remains closed in the resting state. However, as observed in our crystal structure of ligand-bound CusA, the periplasmic cleft is open in the presence of the substrate, which closely coordinates the three-methionine binding site and reveals it to the cytoplasm. Thus, CusA should also be able to transport metal ions from the periplasm, directly through this cleft. Upon entering the central pore of the channel, the electrostatic gradient formed by the inner surfaces of the hexameric CusB and trimeric CusC channels may draw the ions across the outer membrane for final extrusion. The availability of the three-dimensional structures of these efflux transporters and their accessory proteins allows us to determine their mechanism of action and, therefore, should allow us to block their function. Potentially, by the rational design of inhibitors, as demonstrated by AcrB and DARPin. Therefore, the long-term goal of this project is to determine a high-resolution three dimensional structure of a completely assembled tripartite efflux system.

In *Mycobacterium tuberculosis*, pathogenicity and antimicrobial resistance has been shown to depend on the MmpL family of proteins, which is a subset of the RND transporter family. The MmpL transporters have been proposed to transport fatty acid components of the mycobacterial cell wall. In Chapter III, we present the high-resolution crystal structures of two transcriptional regulators of *M. tuberculosis* H37Rv: Rv1816 and Rv3249c. Surprisingly, the electron density maps revealed that the C-terminal ligand binding domains of each protein contained fortuitous ligands. Using GC-MS we identified possible ligands in the protein sample which also fit the unknown densities: palmitic acid in Rv3249c and lauric acid in Rv1816. Using data made available by the TB Systems Biology Consortium, we identified potential DNA binding sites in the TB genome which comprise the *mmpL* genes.

We quantified the binding of Rv1816 and Rv3249c to these sequences and demonstrated that palmitic acid can trigger a conformational change in Rv3249c that disengages it from the DNA.

While we have demonstrated the probability that these proteins regulate the expression of several MmpL transporters, we have not shown their direct effect. Thus, future work should focus on measuring the effect of Rv1816 or Rv3249c on transcription levels of MmpL proteins; possibly by measuring steady state mRNA levels in *M. tuberculosis* or *M. smegmatis*. That only one lauric acid molecule was found in Rv1816 and the binding cavity of one subunit was left empty is intriguing. Although we were not able to determine the apo structure of Rv1816, lauric acid did not seem to trigger a conformational change in our ligand-bound structure. Similarly, lauric acid did not release Rv1816 from its operator DNA in gel-shift assays. Thus, lauric acid occupies the Rv1816 binding cavity like an inducing ligand, but has no effect on its action—it is possible that we stumbled upon a natural inhibitor. Considerable effort should be invested in determining the possible toxicity of lauric acid, and other short-chain fatty acids, on *M. tuberculosis* or *M. smegmatis*, as well as the link between these molecules and other genes that Rv1816 and Rv3249c are proposed to regulate.

In *N. gonorrhoeae*, the potent MtrCDE tripartite efflux pump mediates resistance to diverse hydrophobic antimicrobial agents, such as antibiotics, nonionic detergents, antibacterial peptides, bile salts, and gonadal steroidal hormones. We have previously determined the crystal structures of MtrD and MtrE. In Chapter IV, we present a low-resolution crystal structure of MtrCD. As antibiotic resistance in *N. gonorrhoeae* is

developing into an urgent global problem, we hope to leverage our structural information into the first high-resolution model of any tripartite multidrug efflux pump.

Based on our MtrD structure, our MtrCD structure, and homology with our previous structure of CusBA, we designed chimeric protein constructs that fused MtrC to MtrD in the same protein chain. Ideally, by forcing these proteins in close proximity, the likelihood of the complex forming increases dramatically. Although we were able to obtain minimal levels of expression of chimeric MtrCD in *E. coli*, more effort should be made to express the mg levels of protein necessary for crystallization trials; possibly in *Pichia pastoris* or higher organisms.

In our efforts to assemble the tripartite MtrCDE complex, we discovered that the N-terminal signal peptide and lipidation site of MtrC is not necessary for the drug efflux activity of the pump. Surprisingly, the MIC of *E. coli* expressing the soluble truncated MtrCΔNT, along with MtrD and MtrE, exhibit resistance to TX-100 more than two orders of magnitude above that of *E. coli* missing these three proteins. Further, the expression of full-length MtrC (MtrC FL) or MtrC containing a signal peptide from the homologous membrane fusion protein AcrA (AcrASP-MtrC), actually inhibited the resistance of *E. coli* to TX-100, failing to restore it to even the control level.

As we purified MtrC FL in a large scale and subjected the sample to gel-filtration experiments, we do expect that this protein is well-folded and active. We also found that inclusion of the first 33 residues of MtrC anchored the protein to the membrane of *E. coli*. However, the question remains if MtrC FL or AcrASP-MtrC were targeted to the periplasm, as intended, or if they were attached to the correct membrane. Further experiments are needed to localize these proteins, perhaps by separation and isolation of *E. coli* cell fractions

expressing these proteins or *in vivo* fluorescence experiments. In any case, we can conclude that removal of MtrC from the tripartite pump has a dramatic effect on its function. Thus, the periplasmic adaptor protein MtrC represents an attractive drug target to disrupt antibiotic resistance conferred by the MtrCDE system. Our lab is currently focused on developing peptides which bind specifically to MtrC to prevent its interaction with either MtrD or MtrE.

Finally, in Chapter V, we demonstrate the application of X-ray crystallography to the distinct problem of chloroplast cell division in *Arabidopsis thaliana.* To this end, we have determined the first crystal structures of the PARC6 protein in the intermembrane space (IMS) region of the chloroplast. We also co-crystallized this protein with PDV1, providing the first structural characterization of this interaction that is essential to chloroplast cell division. In our crystal structure, the carboxyl group of PDV1 G272 is located within hydrogen bonding distance of PARC6 W700, revealing the first evidence that PARC6 W700 is essential to the PARC6-PDV1 interaction. Unfortunately, we were not able to carry out mutagenesis experiments to test this, as all attempted mutants of W700 resulted in unstable PARC6 protein.

The arrangement of protein in the asymmetric unit of our initial PARC6 crystal suggests that the protein forms a homodimer. This interaction was also predicted by the program PDBePISA (data not shown) and we observed an identical arrangement in our PARC6-PDV1 co-crystal structure. Surprisingly, a bound peptide was found within the α/β cavity of PARC6 in both structures. In the PARC6-PDV1 crystal, the peptide was, predictably, the C-terminal residues of PDV1. However, the bound ligand in the initial PARC6 structure was identified as a short α-helical segment of PARC6 itself, which was cleaved during pre-crystallization trypsin digestion. It appears that the binding of a peptide

in the α/β cavity of PARC6 induces dimerization, and this result was confirmed by gel-filtration experiments. However, we do not know if the dimerization of PARC6 that we observe has any biological significance. This important question must be addressed in future studies of the roles of PARC6 and PDV1 in chloroplast cell division.